

CHAPTER 1

Section 1.1

1.
 - a. Houston Chronicle, Des Moines Register, Chicago Tribune, Washington Post
 - b. Capital One, Campbell Soup, Merrill Lynch, Pulitzer
 - c. Bill Jasper, Kay Reinke, Helen Ford, David Menedez
 - d. 1.78, 2.44, 3.5, 3.04
2.
 - a. 29.1 yd., 28.3 yd., 24.7 yd., 31.0 yd.
 - b. 432, 196, 184, 321
 - c. 2.1, 4.0, 3.2, 6.3
 - d. 0.07 g, 1.58 g, 7.1 g, 27.2 g
3.
 - a. In a sample of 100 VCRs, what are the chances that more than 20 need service while under warranty? What are the chances that none need service while still under warranty?
 - b. What proportion of all VCRs of this brand and model will need service within the warranty period?

VISIT...

LANZAROTE
Caliente.COM

Chapter 1: Overview and Descriptive Statistics

4.
 - a. Concrete: All living U.S. Citizens, all mutual funds marketed in the U.S., all books published in 1980.
Hypothetical: All grade point averages for University of California undergraduates during the next academic year. Page lengths for all books published during the next calendar year. Batting averages for all major league players during the next baseball season.
 - b. Concrete: Probability: In a sample of 5 mutual funds, what is the chance that all 5 have rates of return which exceeded 10% last year?
Statistics: If previous year rates-of-return for 5 mutual funds were 9.6, 14.5, 8.3, 9.9 and 10.2, can we conclude that the average rate for all funds was below 10%?
Conceptual: Probability: In a sample of 10 books to be published next year, how likely is it that the average number of pages for the 10 is between 200 and 250?
Statistics: If the sample average number of pages for 10 books is 227, can we be highly confident that the average for all books is between 200 and 245?
5.
 - a. No, the relevant conceptual population is all scores of all students who participate in the SI in conjunction with this particular statistics course.
 - b. The advantage to randomly choosing students to participate in the two groups is that we are more likely to get a sample representative of the population at large. If it were left to students to choose, there may be a division of abilities in the two groups which could unnecessarily affect the outcome of the experiment.
 - c. If all students were put in the treatment group there would be no results with which to compare the treatments.
6. One could take a simple random sample of students from all students in the California State University system and ask each student in the sample to report the distance from their hometown to campus. Alternatively, the sample could be generated by taking a stratified random sample by taking a simple random sample from each of the 23 campuses and again asking each student in the sample to report the distance from their hometown to campus. Certain problems might arise with self reporting of distances, such as recording error or poor recall. This study is enumerative because there exists a finite, identifiable population of objects from which to sample.
7. One could generate a simple random sample of all single family homes in the city or a stratified random sample by taking a simple random sample from each of the 10 district neighborhoods. From each of the homes in the sample the necessary variables would be collected. This would be an enumerative study because there exists a finite, identifiable population of objects from which to sample.

Chapter 1: Overview and Descriptive Statistics

- 8.
- a. Number observations equal $2 \times 2 \times 2 = 8$
 - b. This could be called an analytic study because the data would be collected on an existing process. There is no sampling frame.
- 9.
- a. There could be several explanations for the variability of the measurements. Among them could be measuring error, (due to mechanical or technical changes across measurements), recording error, differences in weather conditions at time of measurements, etc.
 - b. This could be called an analytic study because there is no sampling frame.

Section 1.2

- 10.
- a. Minitab generates the following stem-and-leaf display of this data:

```
5|9
6|33588
7|00234677889
8|127
9|077      stem: ones
10|7       leaf: tenths
11|368
```

What constitutes large or small variation usually depends on the application at hand, but an often-used rule of thumb is: the variation tends to be large whenever the spread of the data (the difference between the largest and smallest observations) is large compared to a representative value. Here, 'large' means that the percentage is closer to 100% than it is to 0%. For this data, the spread is $11 - 5 = 6$, which constitutes $6/8 = .75$, or, 75%, of the typical data value of 8. Most researchers would call this a large amount of variation.

- b. The data display is not perfectly symmetric around some middle/representative value. There tends to be some positive skewness in this data.
- c. In Chapter 1, outliers are data points that appear to be *very* different from the pack. Looking at the stem-and-leaf display in part (a), there appear to be no outliers in this data. (Chapter 2 gives a more precise definition of what constitutes an outlier).
- d. From the stem-and-leaf display in part (a), there are 4 values greater than 10. Therefore, the proportion of data values that exceed 10 is $4/27 = .148$, or, about 15%.

Chapter 1: Overview and Descriptive Statistics

11.

6l	034	
6h	667899	
7l	00122244	
7h		Stem=Tens
8l	001111122344	Leaf=Ones
8h	5557899	
9l	03	
9h	58	

This display brings out the gap in the data:
There are no scores in the high 70's.

12. One method of denoting the pairs of stems having equal values is to denote the first stem by L, for 'low', and the second stem by H, for 'high'. Using this notation, the stem-and-leaf display would appear as follows:

3L	1	
3H	56678	
4L	000112222234	
4H	5667888	
5L	144	
5H	58	stem: tenths
6L	2	leaf: hundredths
6H	6678	
7L		
7H	5	

The stem-and-leaf display on the previous page shows that .45 is a good representative value for the data. In addition, the display is not symmetric and appears to be positively skewed. The spread of the data is $.75 - .31 = .44$, which is $.44/.45 = .978$, or about 98% of the typical value of .45. This constitutes a reasonably large amount of variation in the data. The data value .75 is a possible outlier

Chapter 1: Overview and Descriptive Statistics

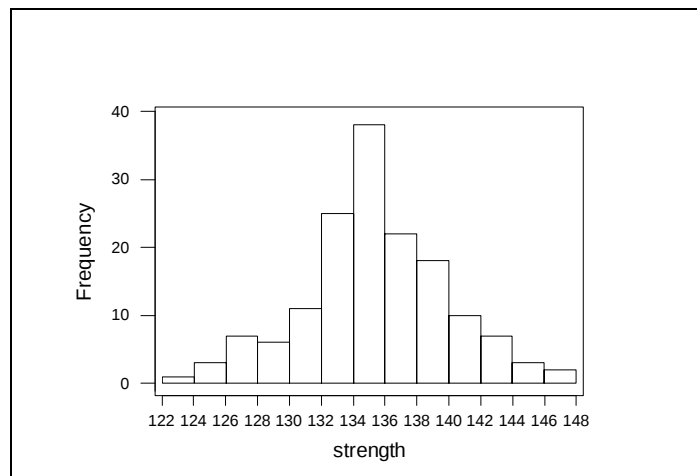
13.

a.

12	2	Leaf = ones
12	445	Stem = tens
12	6667777	
12	889999	
13	00011111111	
13	22222222233333333333333333	
13	44444444444444444445555555555555555555555555555	
13	666666666666666777777777777	
13	88888888888888999999	
14	0000001111	
14	2333333	
14	444	
14	77	

The observations are highly concentrated at 134 – 135, where the display suggests the typical value falls.

b.



The histogram is symmetric and unimodal, with the point of symmetry at approximately 135.

Chapter 1: Overview and Descriptive Statistics

14.

a.

2	23	stem units: 1.0
3	2344567789	leaf units: .10
4	01356889	
5	00001114455666789	
6	0000122223344456667789999	
7	00012233455555668	
8	02233448	
9	012233335666788	
10	2344455688	
11	2335999	
12	37	
13	8	
14	36	
15	0035	
16		
17		
18	9	

- b. A representative value could be the median, 7.0.
- c. The data appear to be highly concentrated, except for a few values on the positive side.
- d. No, the data is skewed to the right, or positively skewed.
- e. The value 18.9 appears to be an outlier, being more than two stem units from the previous value.

15.

Crunchy		Creamy
	2	2
644	3	69
77220	4	145
6320	5	3666
222	6	258
55	7	
0	8	

Both sets of scores are reasonably spread out. There appear to be no outliers. The three highest scores are for the crunchy peanut butter, the three lowest for the creamy peanut butter.

Chapter 1: Overview and Descriptive Statistics

16.

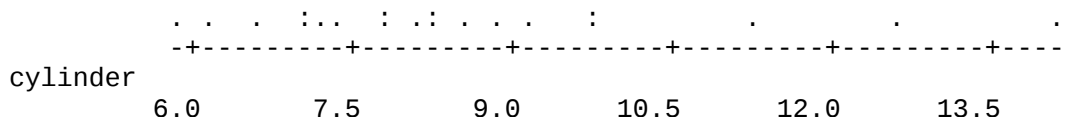
a.

beams		cylinders
9	5	8
88533	6	16
98877643200	7	012488
721	8	13359
770	9	278
7	10	
863	11	2
	12	6
	13	
	14	1

The data appears to be slightly skewed to the right, or positively skewed. The value of 14.1 appears to be an outlier. Three out of the twenty, $3/20$ or $.15$ of the observations exceed 10 Mpa.

- b. The majority of observations are between 5 and 9 Mpa for both beams and cylinders, with the modal class in the 7 Mpa range. The observations for cylinders are more variable, or spread out, and the maximum value of the cylinder observations is higher.

c. Dot Plot



17.

a.

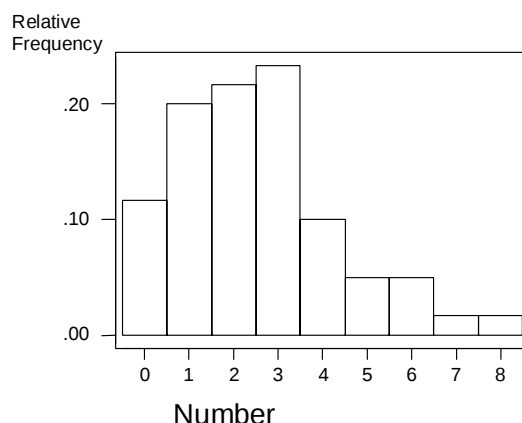
Number	Frequency	RelativeFrequency(Freq/60)
Nonconforming		
0	7	0.117
1	12	0.200
2	13	0.217
3	14	0.233
4	6	0.100
5	3	0.050
6	3	0.050
7	1	0.017
8	1	<u>0.017</u>

doesn't add exactly to 1 because relative frequencies have been rounded 1.001

- b. The number of batches with at most 5 nonconforming items is $7+12+13+14+6+3 = 55$, which is a proportion of $55/60 = .917$. The proportion of batches with (strictly) fewer than 5 nonconforming items is $52/60 = .867$. Notice that these proportions could also have been computed by using the relative frequencies: e.g., proportion of batches with 5 or fewer nonconforming items = $1 - (.05 + .017 + .017) = .916$; proportion of batches with fewer than 5 nonconforming items = $1 - (.05 + .05 + .017 + .017) = .866$.

Chapter 1: Overview and Descriptive Statistics

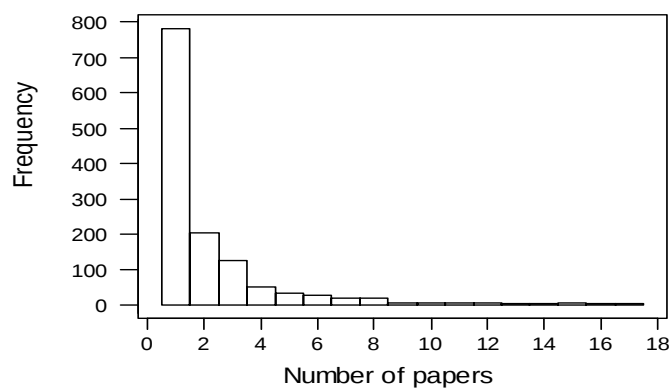
- c. The following is a Minitab histogram of this data. The center of the histogram is somewhere around 2 or 3 and it shows that there is some positive skewness in the data. Using the rule of thumb in Exercise 1, the histogram also shows that there is a lot of spread/variation in this data.



18.

a.

The following histogram was constructed using Minitab:



The most interesting feature of the histogram is the heavy positive skewness of the data.

Note: One way to have Minitab automatically construct a histogram from grouped data such as this is to use Minitab's ability to enter multiple copies of the same number by typing, for example, 784(1) to enter 784 copies of the number 1. The frequency data in this exercise was entered using the following Minitab commands:

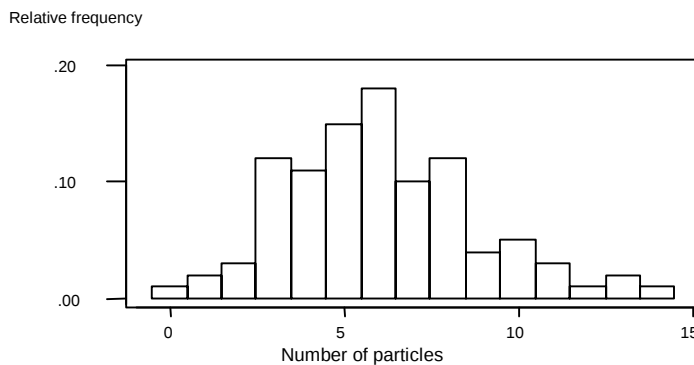
```
MTB > set c1
DATA> 784(1) 204(2) 127(3) 50(4) 33(5) 28(6) 19(7) 19(8)
DATA> 6(9) 7(10) 6(11) 7(12) 4(13) 4(14) 5(15) 3(16) 3(17)
DATA> end
```

Chapter 1: Overview and Descriptive Statistics

- b. From the frequency distribution (or from the histogram), the number of authors who published at least 5 papers is $33+28+19+\dots+5+3+3 = 144$, so the proportion who published 5 or more papers is $144/1309 = .11$, or 11%. Similarly, by adding frequencies and dividing by $n = 1309$, the proportion who published 10 or more papers is $39/1309 = .0298$, or about 3%. The proportion who published more than 10 papers (i.e., 11 or more) is $32/1309 = .0245$, or about 2.5%.
- c. No. Strictly speaking, the class described by ' ≥ 15 ' has no upper boundary, so it is impossible to draw a rectangle above it having finite area (i.e., frequency).
- d. The category 15-17 does have a finite width of 2, so the cumulated frequency of 11 can be plotted as a rectangle of height 6.5 over this interval. The basic rule is to make the area of the bar equal to the class frequency, so $\text{area} = 11 = (\text{width})(\text{height}) = 2(\text{height})$ yields a height of 6.5.

19.

- a. From this frequency distribution, the proportion of wafers that contained at least one particle is $(100-1)/100 = .99$, or 99%. Note that it is much easier to subtract 1 (which is the number of wafers that contain 0 particles) from 100 than it would be to add all the frequencies for 1, 2, 3, ... particles. In a similar fashion, the proportion containing at least 5 particles is $(100 - 1-2-3-12-11)/100 = 71/100 = .71$, or 71%.
- b. The proportion containing between 5 and 10 particles is $(15+18+10+12+4+5)/100 = 64/100 = .64$, or 64%. The proportion that contain strictly between 5 and 10 (meaning strictly *more* than 5 and strictly *less* than 10) is $(18+10+12+4)/100 = 44/100 = .44$, or 44%.
- c. The following histogram was constructed using Minitab. The data was entered using the same technique mentioned in the answer to exercise 8(a). The histogram is *almost* symmetric and unimodal; however, it has a few relative maxima (i.e., modes) and has a very slight positive skew.



Chapter 1: Overview and Descriptive Statistics

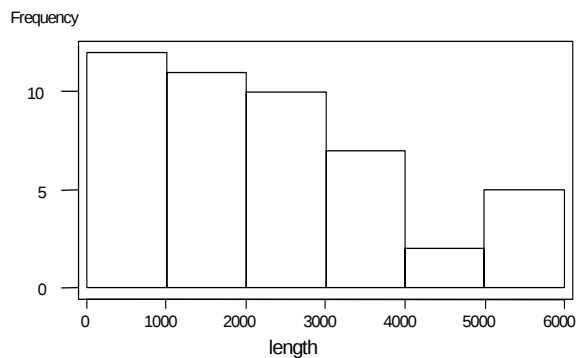
20.

- a. The following stem-and-leaf display was constructed:

0	123334555599	
1	00122234688	stem: thousands
2	1112344477	leaf: hundreds
3	0113338	
4	37	
5	23778	

A typical data value is somewhere in the low 2000's. The display is almost unimodal (the stem at 5 would be considered a mode, the stem at 0 another) and has a positive skew.

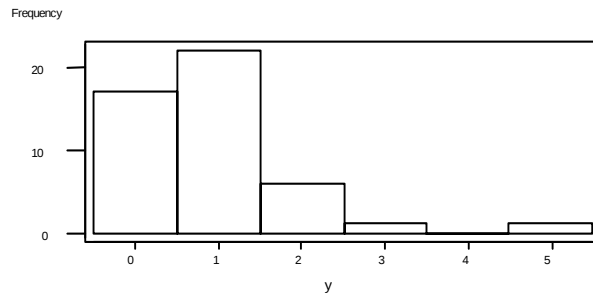
- b. A histogram of this data, using classes of width 1000 centered at 0, 1000, 2000, 6000 is shown below. The proportion of subdivisions with total length less than 2000 is $(12+11)/47 = .489$, or 48.9%. Between 200 and 4000, the proportion is $(7 + 2)/47 = .191$, or 19.1%. The histogram shows the same general shape as depicted by the stem-and-leaf in part (a).



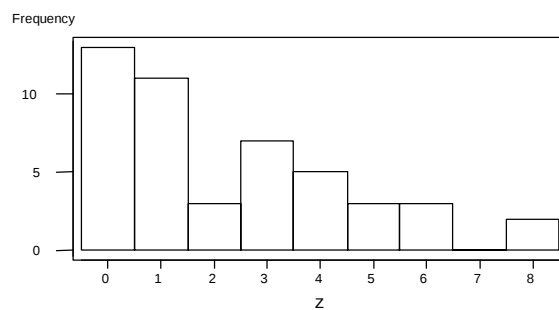
Chapter 1: Overview and Descriptive Statistics

21.

- a. A histogram of the y data appears below. From this histogram, the number of subdivisions having no cul-de-sacs (i.e., $y = 0$) is $17/47 = .362$, or 36.2%. The proportion having at least one cul-de-sac ($y \geq 1$) is $(47-17)/47 = 30/47 = .638$, or 63.8%. Note that subtracting the number of cul-de-sacs with $y = 0$ from the total, 47, is an easy way to find the number of subdivisions with $y \geq 1$.

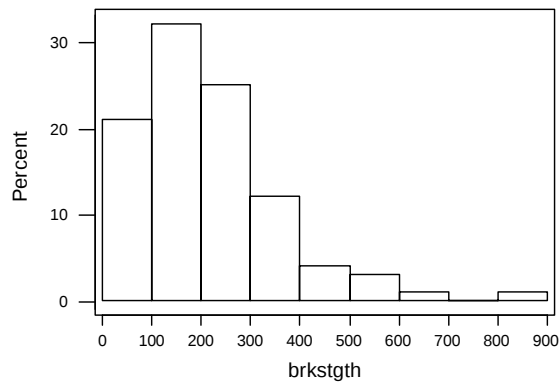


- b. A histogram of the z data appears below. From this histogram, the number of subdivisions with at most 5 intersections (i.e., $z \leq 5$) is $42/47 = .894$, or 89.4%. The proportion having fewer than 5 intersections ($z < 5$) is $39/47 = .830$, or 83.0%.



Chapter 1: Overview and Descriptive Statistics

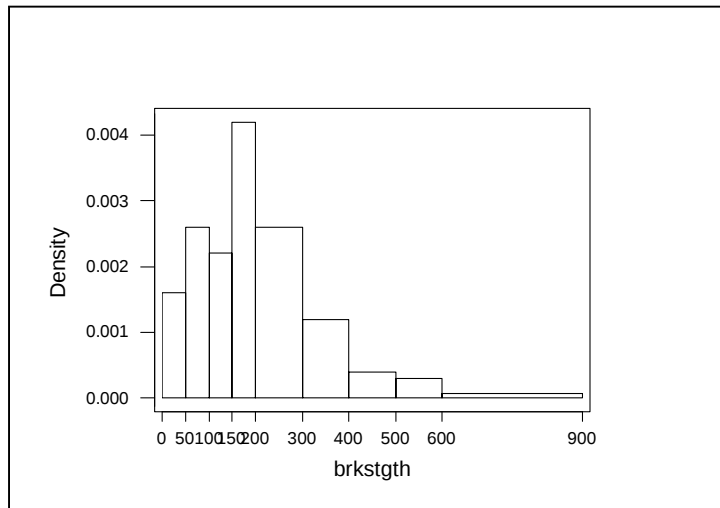
22. A very large percentage of the data values are greater than 0, which indicates that most, but not all, runners do slow down at the end of the race. The histogram is also positively skewed, which means that some runners slow down a *lot* compared to the others. A typical value for this data would be in the neighborhood of 200 seconds. The proportion of the runners who ran the last 5 km faster than they did the first 5 km is very small, about 1% or so.
23. a.



The histogram is skewed right, with a majority of observations between 0 and 300 cycles. The class holding the most observations is between 100 and 200 cycles.

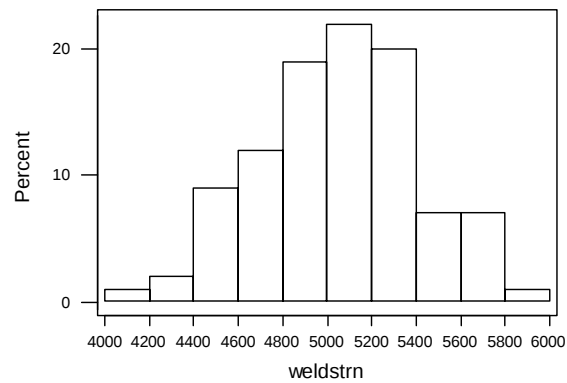
Chapter 1: Overview and Descriptive Statistics

b.



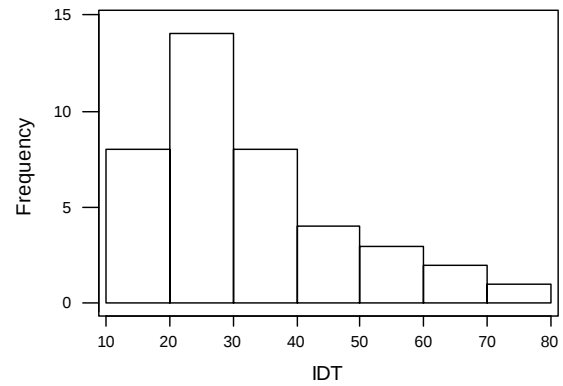
c $[\text{proportion} \geq 100] = 1 - [\text{proportion} < 100] = 1 - .21 = .79$

24.

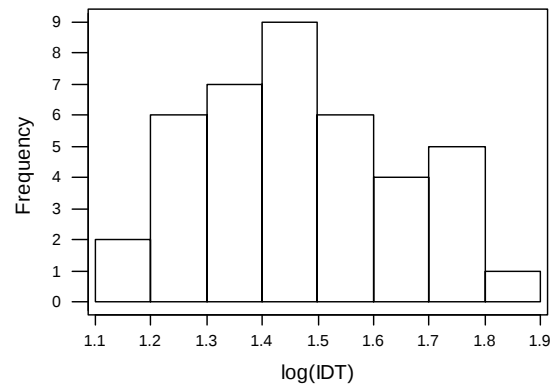


Chapter 1: Overview and Descriptive Statistics

25. Histogram of original data:



Histogram of transformed data:



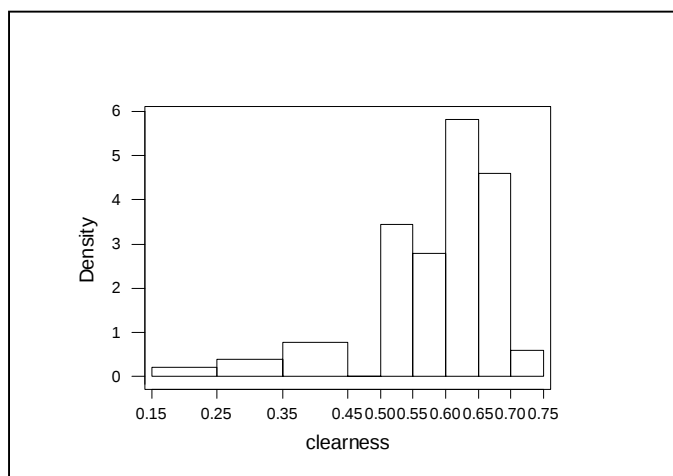
The transformation creates a much more symmetric, mound-shaped histogram.

Chapter 1: Overview and Descriptive Statistics

26.

a.

Class Intervals	Frequency	Rel. Freq.
.15 -< .25	8	0.02192
.25 -< .35	14	0.03836
.35 -< .45	28	0.07671
.45 -< .50	24	0.06575
.50 -< .55	39	0.10685
.55 -< .60	51	0.13973
.60 -< .65	106	0.29041
.65 -< .70	84	0.23014
.70 -< .75	11	0.03014
	n=365	1.00001



- b. The proportion of days with a clearness index smaller than .35 is $\frac{(8 + 14)}{365} = .06$, or 6%.
- c. The proportion of days with a clearness index of at least .65 is $\frac{(84 + 11)}{365} = .26$, or 26%.

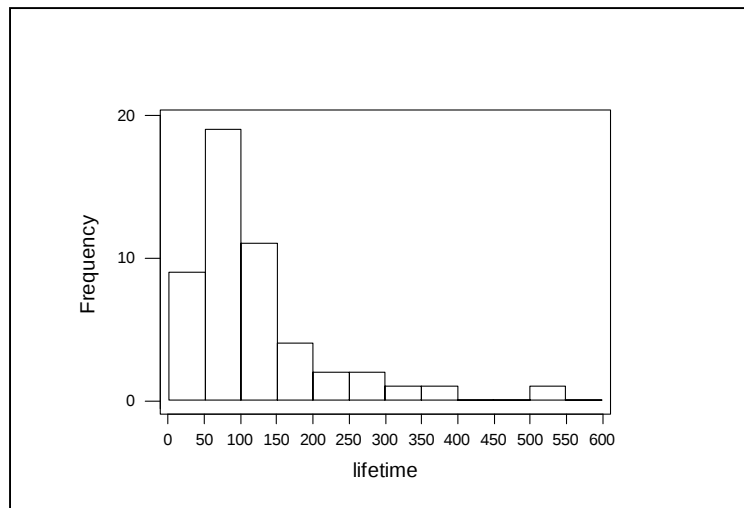
Chapter 1: Overview and Descriptive Statistics

27.

- a. The endpoints of the class intervals overlap. For example, the value 50 falls in both of the intervals '0 – 50' and '50 – 100'.

b.

Class Interval	Frequency	Relative Frequency
0 - < 50	9	0.18
50 - < 100	19	0.38
100 - < 150	11	0.22
150 - < 200	4	0.08
200 - < 250	2	0.04
250 - < 300	2	0.04
300 - < 350	1	0.02
350 - < 400	1	0.02
>= 400	1	0.02
	50	1.00

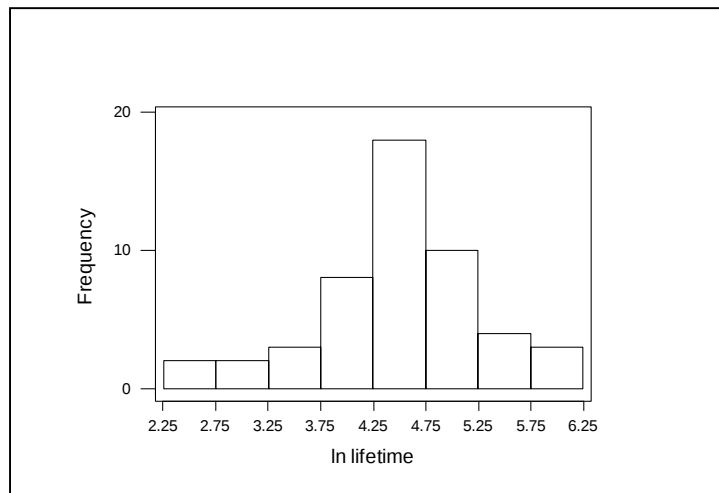


The distribution is skewed to the right, or positively skewed. There is a gap in the histogram, and what appears to be an outlier in the '500 – 550' interval.

Chapter 1: Overview and Descriptive Statistics

c.

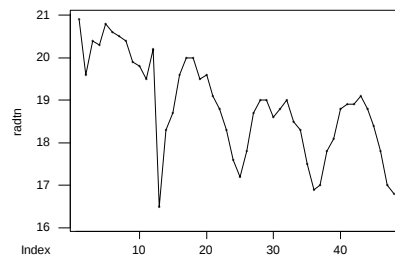
Class Interval	Frequency	Relative Frequency
2.25 - < 2.75	2	0.04
2.75 - < 3.25	2	0.04
3.25 - < 3.75	3	0.06
3.75 - < 4.25	8	0.16
4.25 - < 4.75	18	0.36
4.75 - < 5.25	10	0.20
5.25 - < 5.75	4	0.08
5.75 - < 6.25	3	0.06



The distribution of the natural logs of the original data is much more symmetric than the original.

- d. The proportion of lifetime observations in this sample that are less than 100 is $.18 + .38 = .56$, and the proportion that is at least 200 is $.04 + .04 + .02 + .02 + .02 = .14$.

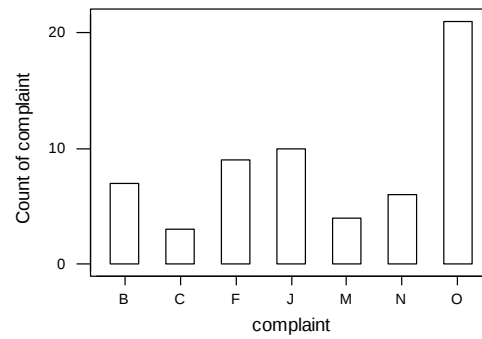
28. There are seasonal trends with lows and highs 12 months apart.



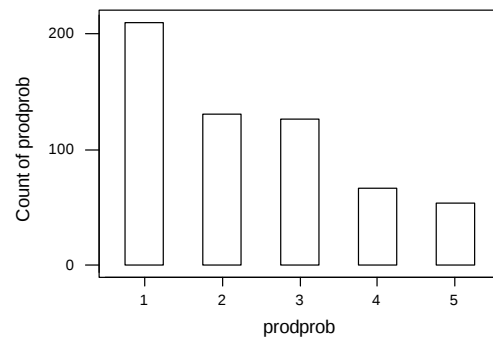
Chapter 1: Overview and Descriptive Statistics

29.

Complaint	Frequency	Relative Frequency
B	7	0.1167
C	3	0.0500
F	9	0.1500
J	10	0.1667
M	4	0.0667
N	6	0.1000
O	21	0.3500
	60	1.0000



30.



1. incorrect component
2. missing component
3. failed component
4. insufficient solder
5. excess solder

Chapter 1: Overview and Descriptive Statistics

31.

Class	Frequency	Relative Frequency	Cumulative Relative Frequency
0.0 - under 4.0	2	2	0.050
4.0 - under 8.0	14	16	0.400
8.0 - under 12.0	11	27	0.675
12.0 - under 16.0	8	35	0.875
16.0 - under 20.0	4	39	0.975
20.0 - under 24.0	0	39	0.975
24.0 - under 28.0	1	40	1.000

32.

- a. The frequency distribution is:

<u>Class</u>	<u>Relative Frequency</u>	<u>Class</u>	<u>Relative Frequency</u>
0-< 150	.193	900-<1050	.019
150-< 300	.183	1050-<1200	.029
300-< 450	.251	1200-<1350	.005
450-< 600	.148	1350-<1500	.004
600-< 750	.097	1500-<1650	.001
750-< 900	.066	1650-<1800	.002
		1800-<1950	.002

The relative frequency distribution is almost unimodal and exhibits a large positive skew. The typical middle value is somewhere between 400 and 450, although the skewness makes it difficult to pinpoint more exactly than this.

- b. The proportion of the fire loads less than 600 is $.193 + .183 + .251 + .148 = .775$. The proportion of loads that are at least 1200 is $.005 + .004 + .001 + .002 + .002 = .014$.
- c. The proportion of loads between 600 and 1200 is $1 - .775 - .014 = .211$.

Chapter 1: Overview and Descriptive Statistics

Section 1.3

33.

- a. $\bar{x} = 192.57$, $\tilde{x} = 189$. The mean is larger than the median, but they are still fairly close together.
- b. Changing the one value, $\bar{x} = 189.71$, $\tilde{x} = 189$. The mean is lowered, the median stays the same.
- c. $\bar{x}_{tr} = 191.0$. $\frac{1}{14} = .07$ or 7% trimmed from each tail.
- d. For $n = 13$, $\Sigma x = (119.7692) \times 13 = 1,557$
For $n = 14$, $\Sigma x = 1,557 + 159 = 1,716$
$$\bar{x} = \frac{1716}{14} = 122.5714 \text{ or } 122.6$$

34.

- a. The sum of the $n = 11$ data points is 514.90, so $\bar{x} = 514.90/11 = 46.81$.
- b. The sample size ($n = 11$) is odd, so there will be a middle value. Sorting from smallest to largest: 4.4 16.4 22.2 30.0 33.1 36.6 40.4 66.7 73.7 81.5 109.9. The sixth value, 36.6 is the middle, or median, value. The mean differs from the median because the largest sample observations are much further from the median than are the smallest values.
- c. Deleting the smallest ($x = 4.4$) and largest ($x = 109.9$) values, the sum of the remaining 9 observations is 400.6. The trimmed mean $\bar{x}_{tr}$ is $400.6/9 = 44.51$. The trimming percentage is $100(1/11) \approx 9.1\%$. $\bar{x}_{tr}$ lies between the mean and median.

35.

- a. The sample mean is $\bar{x} = (100.4/8) = 12.55$.

The sample size ($n = 8$) is even. Therefore, the sample median is the average of the $(n/2)$ and $(n/2) + 1$ values. By sorting the 8 values in order, from smallest to largest: 8.0 8.9 11.0 12.0 13.0 14.5 15.0 18.0, the fourth and fifth values are 12 and 13. The sample median is $(12.0 + 13.0)/2 = 12.5$.

The 12.5% trimmed mean requires that we first trim $(.125)(n)$ or 1 value from the ends of the ordered data set. Then we average the remaining 6 values. The 12.5% trimmed mean $\bar{x}_{tr(12.5)}$ is $74.4/6 = 12.4$.

All three measures of center are similar, indicating little skewness to the data set.

- b. The smallest value (8.0) could be increased to any number below 12.0 (a change of less than 4.0) without affecting the value of the sample median.

Chapter 1: Overview and Descriptive Statistics

- c. The values obtained in part (a) can be used directly. For example, the sample mean of 12.55 psi could be re-expressed as

$$(12.55 \text{ psi}) \times \left(\frac{1 \text{ ksi}}{2.2 \text{ psi}} \right) = 5.70 \text{ ksi}.$$

36.

- a. A stem-and leaf display of this data appears below:

32	55	stem: ones
33	49	leaf: tenths
34		
35	6699	
36	34469	
37	03345	
38	9	
39	2347	
40	23	
41		
42	4	

The display is reasonably symmetric, so the mean and median will be close.

- b. The sample mean is $\bar{x} = 9638/26 = 370.7$. The sample median is $\tilde{x} = (369+370)/2 = 369.50$.
- c. The largest value (currently 424) could be increased by any amount. Doing so will not change the fact that the middle two observations are 369 and 170, and hence, the median will not change. However, the value $x = 424$ can not be changed to a number less than 370 (a change of $424-370 = 54$) since that *will* lower the values(s) of the two middle observations.
- d. Expressed in minutes, the mean is $(370.7 \text{ sec})/(60 \text{ sec}) = 6.18 \text{ min}$; the median is 6.16 min.

37. $\bar{x} = 12.01$, $\tilde{x} = 11.35$, $\bar{x}_{tr(10)} = 11.46$. The median or the trimmed mean would be good choices because of the outlier 21.9.

38.

- a. The reported values are (in increasing order) 110, 115, 120, 120, 125, 130, 130, 135, and 140. Thus the median of the reported values is 125.
- b. 127.6 is reported as 130, so the median is now 130, a very substantial change. When there is rounding or grouping, the median can be highly sensitive to small change.

Chapter 1: Overview and Descriptive Statistics

39.

a. $\Sigma x_i = 16.475$ so $\bar{x} = \frac{16.475}{16} = 1.0297$

$$\tilde{x} = \frac{(1.007 + 1.011)}{2} = 1.009$$

- b. 1.394 can be decreased until it reaches 1.011 (the largest of the 2 middle values) – i.e. by $1.394 - 1.011 = .383$. If it is decreased by more than .383, the median will change.

40.

$$\tilde{x} = 60.8$$

$$\bar{x}_{tr(25)} = 59.3083$$

$$\bar{x}_{tr(10)} = 58.3475$$

$$\bar{x} = 58.54$$

All four measures of center have about the same value.

41.

a. $\frac{7}{10} = .70$

b. $\bar{x} = .70$ = proportion of successes

c. $\frac{s}{25} = .80$ so $s = (0.80)(25) = 20$

total of 20 successes

$20 - 7 = 13$ of the new cars would have to be successes

42.

a. $\bar{y} = \frac{\Sigma y_i}{n} = \frac{\Sigma(x_i + c)}{n} = \frac{\Sigma x_i}{n} + \frac{nc}{n} = \bar{x} + c$

$\tilde{y}$ = the median of $(x_1 + c, x_2 + c, \dots, x_n + c)$ = median of

$$(x_1, x_2, \dots, x_n) + c = \tilde{x} + c$$

b. $\bar{y} = \frac{\Sigma y_i}{n} = \frac{\Sigma(x_i \cdot c)}{n} = \frac{c \Sigma x_i}{n} = c\bar{x}$

$$\tilde{y} = (cx_1, cx_2, \dots, cx_n) = c \cdot \text{median}(x_1, x_2, \dots, x_n) = c\tilde{x}$$

43. median = $\frac{(57 + 79)}{2} = 68.0$, 20% trimmed mean = 66.2, 30% trimmed mean = 67.5.

Section 1.4

44.

a. range = $49.3 - 23.5 = 25.8$

b.

x_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	x_i^2
29.5	-1.53	2.3409	870.25
49.3	18.27	333.7929	2430.49
30.6	-0.43	0.1849	936.36
28.2	-2.83	8.0089	795.24
28.0	-3.03	9.1809	784.00
26.3	-4.73	22.3729	691.69
33.9	2.87	8.2369	1149.21
29.4	-1.63	2.6569	864.36
23.5	-7.53	56.7009	552.25
31.6	0.57	0.3249	998.56
$\Sigma x = 310.3$	$\Sigma(x_i - \bar{x}) = 0$	$\Sigma(x_i - \bar{x})^2 = 443.801$	$\Sigma(x_i^2) = 10,072.41$

$$\bar{x} = 31.03$$

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} = \frac{443.801}{9} = 49.3112$$

c. $s = \sqrt{s^2} = 7.0222$

d. $s^2 = \frac{\Sigma x^2 - (\Sigma x)^2 / n}{n-1} = \frac{10,072.41 - (310.3)^2 / 10}{9} = 49.3112$

45.

a. $\bar{x} = \frac{1}{n} \sum x_i = 577.9/5 = 115.58$. Deviations from the mean:

$116.4 - 115.58 = .82$, $115.9 - 115.58 = .32$, $114.6 - 115.58 = -.98$,
 $115.2 - 115.58 = -.38$, and $115.8 - 115.58 = .22$.

b. $s^2 = [(.82)^2 + (.32)^2 + (-.98)^2 + (-.38)^2 + (.22)^2] / (5-1) = 1.928/4 = .482$,
 so $s = .694$.

c. $\sum x_i^2 = 66,795.61$, so $s^2 = \frac{1}{n-1} \left[\sum x_i^2 - \frac{1}{n} \left(\sum x_i \right)^2 \right] =$

$[66,795.61 - (577.9)^2 / 5] / 4 = 1.928/4 = .482$.

d. Subtracting 100 from all values gives $\bar{x} = 15.58$, all deviations are the same as in part b, and the transformed variance is identical to that of part b.

Chapter 1: Overview and Descriptive Statistics

46.

a. $\bar{x} = \frac{1}{n} \sum_i x_i = 14438/5 = 2887.6$. The sorted data is: 2781 2856 2888 2900 3013,
so the sample median is $\tilde{x} = 2888$.

b. Subtracting a constant from each observation shifts the data, but does not change its sample variance (Exercise 16). For example, by subtracting 2700 from each observation we get the values 81, 200, 313, 156, and 188, which are smaller (fewer digits) and easier to work with. The sum of squares of this transformed data is 204210 and its sum is 938, so the computational formula for the variance gives $s^2 = [204210 - (938)^2/5]/(5-1) = 7060.3$.

47. The sample mean, $\bar{x} = \frac{1}{n} \sum x_i = \frac{1}{10}(1,162) = \bar{x} = 116.2$.

$$\text{The sample standard deviation, } s = \sqrt{\frac{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}{n-1}} = \sqrt{\frac{140,992 - \frac{(1,162)^2}{10}}{9}} = 25.75$$

On average, we would expect a fracture strength of 116.2. In general, the size of a typical deviation from the sample mean (116.2) is about 25.75. Some observations may deviate from 116.2 by more than this and some by less.

48. Using the computational formula, $s^2 = \frac{1}{n-1} \left[\sum_i x_i^2 - \frac{1}{n} \left(\sum_i x_i \right)^2 \right] =$

$[3,587,566 - (9638)^2/26]/(26-1) = 593.3415$, so $s = 24.36$. In general, the size of a typical deviation from the sample mean (370.7) is about 24.4. Some observations may deviate from 370.7 by a little more than this, some by less.

49.

a. $\Sigma x = 2.75 + \dots + 3.01 = 56.80$, $\Sigma x^2 = (2.75)^2 + \dots + (3.01)^2 = 197.8040$

b. $s^2 = \frac{197.8040 - (56.80)^2/17}{16} = \frac{8.0252}{16} = .5016$, $s = .708$

Chapter 1: Overview and Descriptive Statistics

50. First, we need $\bar{x} = \frac{1}{n} \sum x_i = \frac{1}{27} (20,179) = 747.37$. Then we need the sample standard

deviation $s = \sqrt{\frac{24,657,511 - \frac{(20,179)^2}{27}}{26}} = 606.89$. The maximum award should be $\bar{x} + 2s = 747.37 + 2(606.89) = 1961.16$, or in dollar units, \$1,961,160. This is quite a bit less than the \$3.5 million that was awarded originally.

51.

- a. $\Sigma x = 2563$ and $\Sigma x^2 = 368,501$, so

$$s^2 = \frac{[368,501 - (2563)^2 / 19]}{18} = 1264.766 \text{ and } s = 35.564$$

- b. If y = time in minutes, then $y = cx$ where $c = \frac{1}{60}$, so

$$s_y^2 = c^2 s_x^2 = \frac{1264.766}{3600} = .351 \text{ and } s_y = cs_x = \frac{35.564}{60} = .593$$

52. Let d denote the fifth deviation. Then $.3 + .9 + 1.0 + 1.3 + d = 0$ or $3.5 + d = 0$, so $d = -3.5$. One sample for which these are the deviations is $x_1 = 3.8$, $x_2 = 4.4$, $x_3 = 4.5$, $x_4 = 4.8$, $x_5 = 0$. (obtained by adding 3.5 to each deviation; adding any other number will produce a different sample with the desired property)

53.

- a. lower half: 2.34 2.43 2.62 2.74 2.74 2.75 2.78 3.01 3.46
upper half: 3.46 3.56 3.65 3.85 3.88 3.93 4.21 4.33 4.52
Thus the lower fourth is 2.74 and the upper fourth is 3.88.

- b. $f_s = 3.88 - 2.74 = 1.14$

- c. f_s wouldn't change, since increasing the two largest values does not affect the upper fourth.
- d. By at most .40 (that is, to anything not exceeding 2.74), since then it will not change the lower fourth.
- e. Since n is now even, the lower half consists of the smallest 9 observations and the upper half consists of the largest 9. With the lower fourth = 2.74 and the upper fourth = 3.93, $f_s = 1.19$.

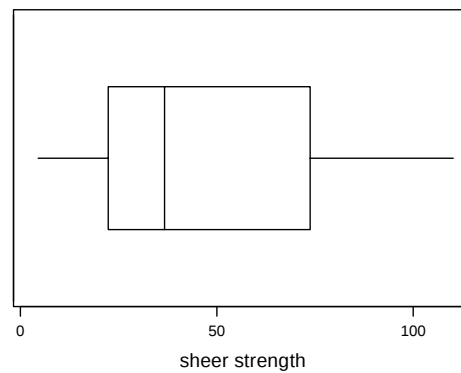
Chapter 1: Overview and Descriptive Statistics

54.

- a. The lower half of the data set: 4.4 16.4 22.2 30.0 33.1 36.6, whose median, and therefore, the lower quartile, is $\frac{(22.2 + 30.0)}{2} = 26.1$.
The top half of the data set: 36.6 40.4 66.7 73.7 81.5 109.9, whose median, and therefore, the upper quartile, is $\frac{(66.7 + 73.7)}{2} = 70.2$.
So, the IQR = $(70.2 - 26.1) = 44.1$

b.

A boxplot (created in Minitab) of this data appears below:



There is a slight positive skew to the data. The variation seems quite large. There are no outliers.

- c. An observation would need to be further than $1.5(44.1) = 66.15$ units below the lower quartile $[(26.1 - 66.15) = -40.05 \text{ units}]$ or above the upper quartile $[(70.2 + 66.15) = 136.35 \text{ units}]$ to be classified as a mild outlier. Notice that, in this case, an outlier on the lower side would not be possible since the shear strength variable cannot have a negative value.

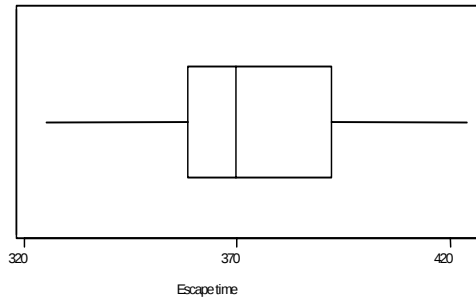
An extreme outlier would fall $(3)(44.1) = 132.3$ or more units below the lower, or above the upper quartile. Since the minimum and maximum observations in the data are 4.4 and 109.9 respectively, we conclude that there are no outliers, of either type, in this data set.

- d. Not until the value $x = 109.9$ is lowered below 73.7 would there be any change in the value of the upper quartile. That is, the value $x = 109.9$ could not be decreased by more than $(109.9 - 73.7) = 36.2$ units.

Chapter 1: Overview and Descriptive Statistics

55.

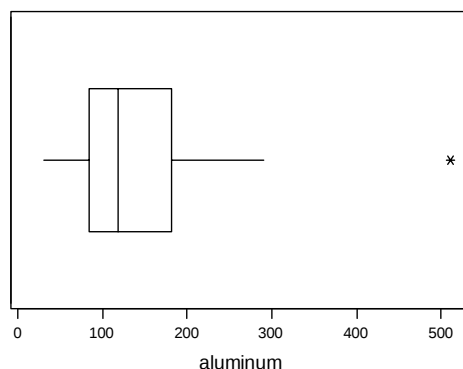
- a. Lower half of the data set: 325 325 334 339 356 356 359 359 363 364 364 366 369, whose median, and therefore the lower quartile, is 359 (the 7th observation in the sorted list).
The top half of the data is 370 373 373 374 375 389 392 393 394 397 402 403 424, whose median, and therefore the upper quartile is 392. So, the $IQR = 392 - 359 = 33$.
- b. $1.5(IQR) = 1.5(33) = 49.5$ and $3(IQR) = 3(33) = 99$. Observations that are further than 49.5 below the lower quartile (i.e., $359 - 49.5 = 309.5$ or less) or more than 49.5 units above the upper quartile (greater than $392 + 49.5 = 441.5$) are classified as 'mild' outliers. 'Extreme' outliers would fall 99 or more units below the lower, or above the upper, quartile. Since the minimum and maximum observations in the data are 325 and 424, we conclude that there are no mild outliers in this data (and therefore, no 'extreme' outliers either).
- c. A boxplot (created by Minitab) of this data appears below. There is a slight positive skew to the data, but it is not far from being symmetric. The variation, however, seems large (the spread $424 - 325 = 99$ is a large percentage of the median/typical value)



- d. Not until the value $x = 424$ is lowered below the upper quartile value of 392 would there be any change in the value of the upper quartile. That is, the value $x = 424$ could not be decreased by more than $424 - 392 = 32$ units.

Chapter 1: Overview and Descriptive Statistics

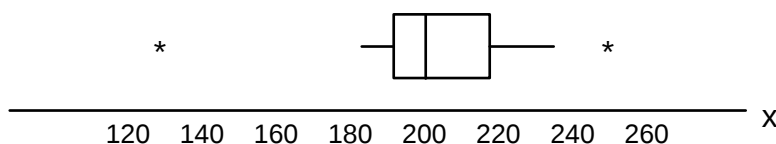
56. A boxplot (created in Minitab) of this data appears below.



There is a slight positive skew to this data. There is one extreme outlier ($x=511$). Even when removing the outlier, the variation is still moderately large.

- 57.

- a. $1.5(IQR) = 1.5(216.8-196.0) = 31.2$ and $3(IQR) = 3(216.8-196.0) = 62.4$.
Mild outliers: observations below $196-31.2 = 164.6$ or above $216.8+31.2 = 248$.
Extreme outliers: observations below $196-62.4 = 133.6$ or above $216.8+62.4 = 279.2$. Of the observations given, 125.8 is an extreme outlier and 250.2 is a mild outlier.
- b. A boxplot of this data appears below. There is a bit of positive skew to the data but, except for the two outliers identified in part (a), the variation in the data is relatively small.



58. The most noticeable feature of the comparative boxplots is that machine 2's sample values have considerably more variation than does machine 1's sample values. However, a typical value, as measured by the median, seems to be about the same for the two machines. The only outlier that exists is from machine 1.

Chapter 1: Overview and Descriptive Statistics

59.

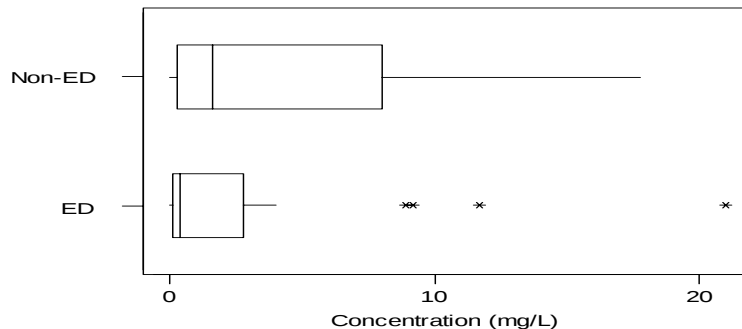
- a. ED: median = .4 (the 14th value in the *sorted* list of data). The lower quartile (median of the lower half of the data, including the median, since n is odd) is $(.1 + .1)/2 = .1$. The upper quartile is $(2.7 + 2.8)/2 = 2.75$. Therefore, $IQR = 2.75 - .1 = 2.65$.

Non-ED: median = $(1.5 + 1.7)/2 = 1.6$. The lower quartile (median of the lower 25 observations) is .3; the upper quartile (median of the upper half of the data) is 7.9. Therefore, $IQR = 7.9 - .3 = 7.6$.

- b. ED: mild outliers are less than $.1 - 1.5(2.65) = -3.875$ or greater than $2.75 + 1.5(2.65) = 6.725$. Extreme outliers are less than $.1 - 3(2.65) = -7.85$ or greater than $2.75 + 3(2.65) = 10.7$. So, the two largest observations (11.7, 21.0) are extreme outliers and the next two largest values (8.9, 9.2) are mild outliers. There are no outliers at the lower end of the data.

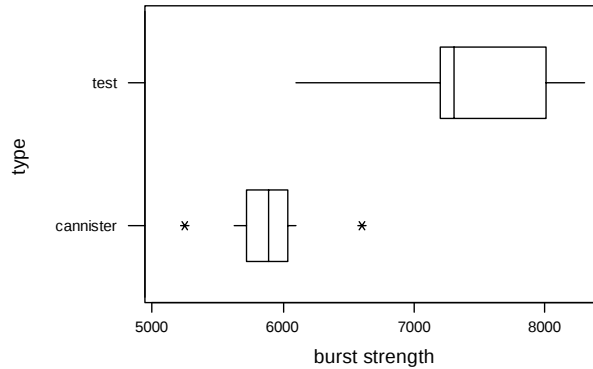
Non-ED: mild outliers are less than $.3 - 1.5(7.6) = -11.1$ or greater than $7.9 + 1.5(7.6) = 19.3$. Note that there are no mild outliers in the data, hence there can not be any extreme outliers either.

- c. A comparative boxplot appears below. The outliers in the ED data are clearly visible. There is noticeable positive skewness in both samples; the Non-Ed data has more variability than the Ed data; the typical values of the ED data tend to be smaller than those for the Non-ED data.



Chapter 1: Overview and Descriptive Statistics

- 60.** A comparative boxplot (created in Minitab) of this data appears below.



The burst strengths for the test nozzle closure welds are quite different from the burst strengths of the production canister nozzle welds.

The test welds have much higher burst strengths and the burst strengths are much more variable.

The production welds have more consistent burst strength and are consistently lower than the test welds. The production welds data does contain 2 outliers.

- 61.** Outliers occur in the 6 a.m. data. The distributions at the other times are fairly symmetric. Variability and the 'typical' values in the data increase a little at the 12 noon and 2 p.m. times.

Supplementary Exercises

62. To somewhat simplify the algebra, begin by subtracting 76,000 from the original data. This transformation will affect each data value and the mean. It will not affect the standard deviation.

$$x_1 = 683, \quad x_2 = 1,048, \quad \bar{y} = 831$$

$$n\bar{x} = (4)(831) = 3,324 \text{ so, } x_1 + x_2 + x_3 + x_4 = 3,324$$

$$\text{and } x_2 + x_3 = 3,324 - x_1 - x_4 = 1,593 \text{ and } x_3 = (1,593 - x_2)$$

$$\text{Next, } s^2 = (180)^2 = \left[\frac{\sum x_i^2 - \frac{(3,324)^2}{4}}{3} \right]$$

$$\text{So, } \sum x_i^2 = 2,859,444, \quad x_1^2 + x_2^2 + x_3^2 + x_4^2 = 2,859,444 \text{ and}$$

$$x_2^2 + x_3^2 = 2,859,444 - x_1^2 - x_4^2 = 1,294,651$$

By substituting $x_3 = (1,593 - x_2)$ we obtain the equation

$$x_2^2 + (1,593 - x_2)^2 - 1,294,651 = 0.$$

$$x_2^2 - 1,593x_2 + 621,499 = 0$$

Evaluating for x_2 we obtain $x_2 = 682.8635$ and $x_3 = 1,593 - 682.8635 = 910.1365$.

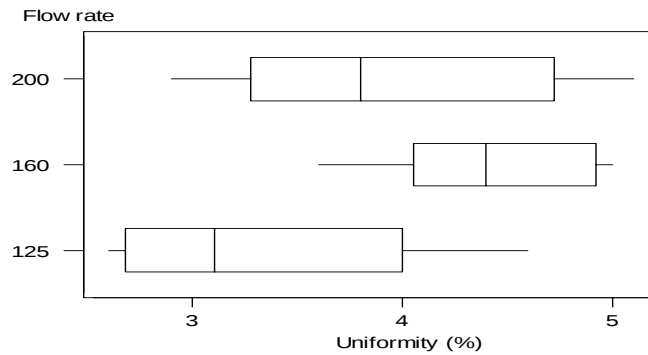
Thus, $x_2 = 76,683$ $x_3 = 76,910$.

Chapter 1: Overview and Descriptive Statistics

63.

Flow rate	Median	Lower quartile	Upper quartile	IQR	1.5(IQR)	3(IQR)
125	3.1	2.7	3.8	1.1	1.65	.3
160	4.4	4.2	4.9	.7	1.05	.1
200	3.8	3.4	4.6	1.2	1.80	3.6

There are no outliers in the three data sets. However, as the comparative boxplot below shows, the three data sets differ with respect to their central values (the medians are different) and the data for flow rate 160 is somewhat less variable than the other data sets. Flow rates 125 and 200 also exhibit a small degree of positive skewness.



Chapter 1: Overview and Descriptive Statistics

64.

6		34	
7		17	stem=ones
8		4589	leaf=tenths
9		1	
10		12667789	
11		122499	
12		2	
13		1	

$$\bar{x} = 9.9556, \tilde{x} = 10.6$$

$$s = 1.7594$$

$$n = 27$$

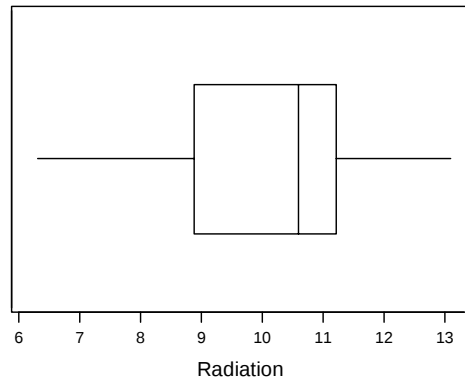
$$\text{lower fourth} = 8.85, \text{upper fourth} = 11.15$$

$$f_s = 2.3$$

$$8.85 - (1.5)(2.3) = 5.4$$

$$11.15 + (1.5)(2.3) = 14.6$$

no outliers



There are no outliers. The distribution is skewed to the left.

Chapter 1: Overview and Descriptive Statistics

65.

a. HC data: $\sum_i x_i^2 = 2618.42$ and $\sum_i x_i = 96.8$,

so $s^2 = [2618.42 - (96.8)^2/4]/3 = 91.953$

and the sample standard deviation is $s = 9.59$.

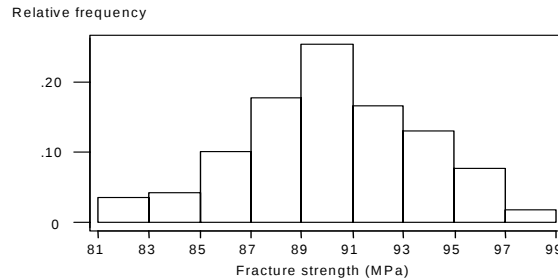
CO data: $\sum_i x_i^2 = 145645$ and $\sum_i x_i = 735$, so $s^2 = [145645 - (735)^2/4]/3 =$

3529.583 and the sample standard deviation is $s = 59.41$.

b. The mean of the HC data is $96.8/4 = 24.2$; the mean of the CO data is $735/4 = 183.75$. Therefore, the coefficient of variation of the HC data is $9.59/24.2 = .3963$, or 39.63%. The coefficient of variation of the CO data is $59.41/183.75 = .3233$, or 32.33%. Thus, even though the CO data has a larger standard deviation than does the HC data, it actually exhibits *less* variability (in percentage terms) around its average than does the HC data.

66.

a. The histogram appears below. A representative value for this data would be $x = 90$. The histogram is reasonably symmetric, unimodal, and somewhat bell-shaped. The variation in the data is not small since the spread of the data ($99-81 = 18$) constitutes about 20% of the typical value of 90.



b. The proportion of the observations that are at least 85 is $1 - (6+7)/169 = .9231$. The proportion less than 95 is $1 - (22+13+3)/169 = .7751$.

c. $x = 90$ is the midpoint of the class $89-<91$, which contains 43 observations (a relative frequency of $43/169 = .2544$). Therefore about half of this frequency, .1272, should be added to the relative frequencies for the classes to the left of $x = 90$. That is, the approximate proportion of observations that are less than 90 is $.0355 + .0414 + .1006 + .1775 + .1272 = .4822$.

67.

$$\begin{aligned}\sum x_i &= 163.2 \\ 100\left(\frac{1}{15}\right)\%trimmedmean &= \frac{163.2 - 8.5 - 15.6}{13} = 10.70 \\ 100\left(\frac{2}{15}\right)\%trimmedmean &= \frac{163.2 - 8.5 - 8.8 - 15.6 - 13.7}{11} = 10.60 \\ \therefore \frac{1}{2}(100)\left(\frac{1}{15}\right) + \frac{1}{2}(100)\left(\frac{2}{15}\right) &= 100\left(\frac{1}{10}\right) = 10\%trimmedmean \\ &= \frac{1}{2}(10.70) + \frac{1}{2}(10.60) = 10.65\end{aligned}$$

68.

a.
$$\frac{d}{dc\left\{\sum (x_i - c)^2\right\}} = \frac{\sum d}{dc(x_i - c)^2} = -2\sum (x_i - c) = 0 \Rightarrow \sum (x_i - c) = 0$$

$$\Rightarrow \sum x_i - \sum c = 0 \Rightarrow \sum x_i - nc = 0 \Rightarrow nc = \sum x_i \Rightarrow c = \frac{\sum x_i}{n} = \bar{x}.$$

b.
$$\sum (x_i - \bar{x})^2 \text{ is smaller than } \sum (x_i - \mu)^2.$$

69.

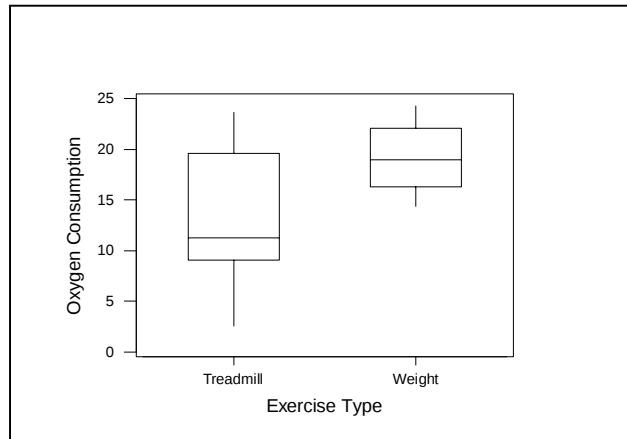
a.
$$\begin{aligned}\bar{y} &= \frac{\sum y_i}{n} = \frac{\sum (ax_i + b)}{n} = \frac{a\sum x_i + b}{n} = a\bar{x} + b. \\ s_y^2 &= \frac{\sum (y_i - \bar{y})^2}{n-1} = \frac{\sum (ax_i + b - (a\bar{x} + b))^2}{n-1} = \frac{\sum (ax_i - a\bar{x})^2}{n-1} \\ &= \frac{a^2 \sum (x_i - \bar{x})^2}{n-1} = a^2 s_x^2.\end{aligned}$$

b.

$$\begin{aligned}x &= ^\circ C, y = ^\circ F \\ \bar{y} &= \frac{9}{5}(87.3) + 32 = 189.14 \\ s_y &= \sqrt{s_y^2} = \sqrt{\left(\frac{9}{5}\right)^2 (1.04)^2} = \sqrt{3.5044} = 1.872\end{aligned}$$

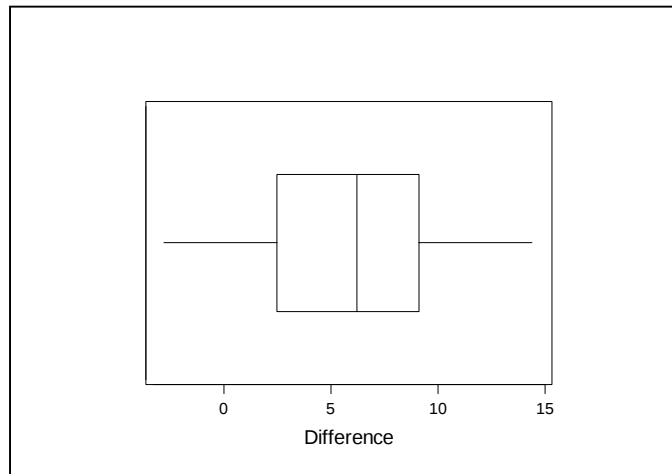
70.

a.



There is a significant difference in the variability of the two samples. The weight training produced much higher oxygen consumption, on average, than the treadmill exercise, with the median consumptions being approximately 20 and 11 liters, respectively.

- b. Subtracting the y from the x for each subject, the differences are 3.3, 9.1, 10.4, 9.1, 6.2, 2.5, 2.2, 8.4, 8.7, 14.4, 2.5, -2.8, -0.4, 5.0, and 11.5.



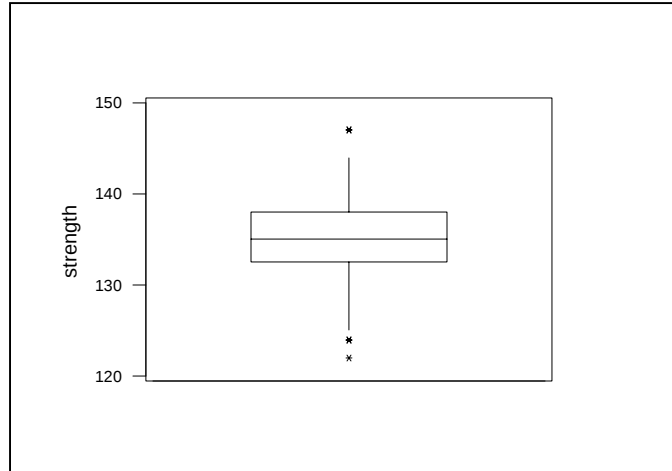
The majority of the differences are positive, which suggests that the weight training produced higher oxygen consumption for most subjects. The median difference is about 6 liters.

Chapter 1: Overview and Descriptive Statistics

71.

- a. The mean, median, and trimmed mean are virtually identical, which suggests symmetry. If there are outliers, they are balanced. The range of values is only 25.5, but half of the values are between 132.95 and 138.25.

b.

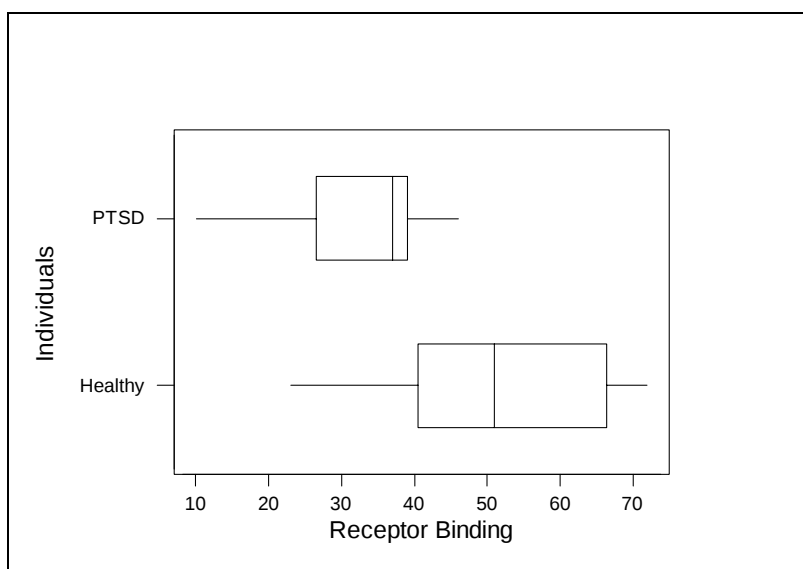


The boxplot also displays the symmetry, and adds a visual of the outliers, two on the lower end, and one on the upper.

Chapter 1: Overview and Descriptive Statistics

72. A table of summary statistics, a stem and leaf display, and a comparative boxplot are below. The healthy individuals have higher receptor binding measure on average than the individuals with PTSD. There is also more variation in the healthy individuals' values. The distribution of values for the healthy is reasonably symmetric, while the distribution for the PTSD individuals is negatively skewed. The box plot indicates that there are no outliers, and confirms the above comments regarding symmetry and skewness.

	PTSD	Healthy				
Mean	32.92	52.23		1	0	stem = tens
Median	37	51	3	2	058	leaf = ones
Std Dev	9.93	14.86	9	3	1578899	
Min	10	23	7310	4	26	
Max	46	72	81	5		
			9763	6		
			2	7		



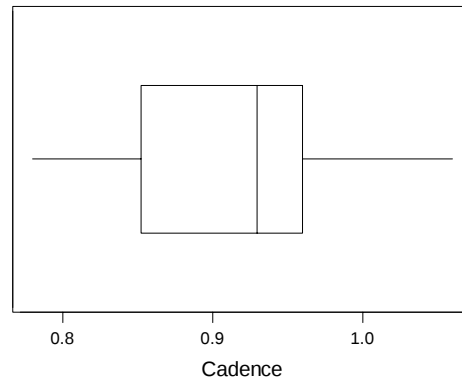
Chapter 1: Overview and Descriptive Statistics

73.

```
0.7|8          stem=tenths
0.8|11556      leaf=hundredths
0.9|2233335566
1.0|0566
```

$$\bar{x} = .9255, s = .0809, \tilde{x} = .93$$

$$lowerfourth = .855, upperfourth = .96$$



The data appears to be a bit skewed toward smaller values (negatively skewed). There are no outliers. The mean and the median are close in value.

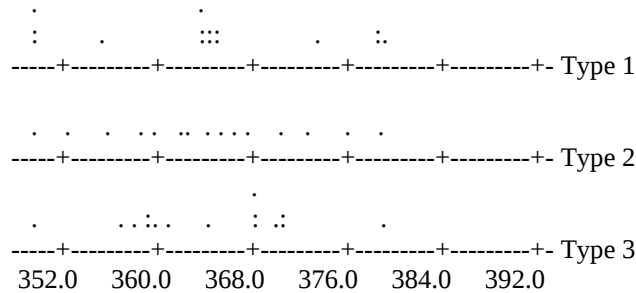
74.

- a. Mode = .93. It occurs four times in the data set.
- b. The Modal Category is the one in which the most observations occur.

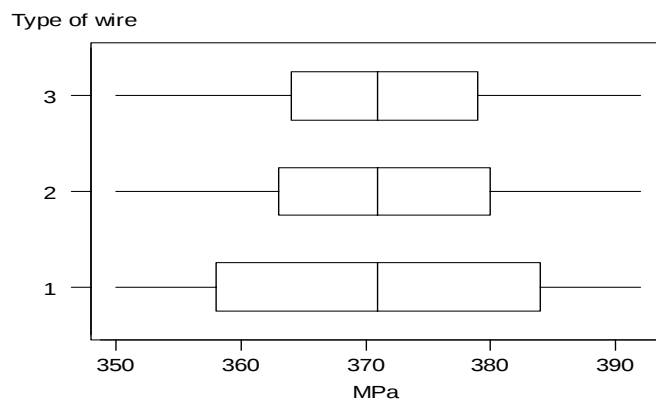
Chapter 1: Overview and Descriptive Statistics

75.

- a. The median is the same (371) in each plot and all three data sets are very symmetric. In addition, all three have the same minimum value (350) and same maximum value (392). Moreover, all three data sets have the same lower (364) and upper quartiles (378). So, all three boxplots will be *identical*.
- b. A comparative dotplot is shown below. These graphs show that there are differences in the variability of the three data sets. They also show differences in the way the values are distributed in the three data sets.



- c. The boxplot in (a) is not capable of detecting the differences among the data sets. The primary reason is that boxplots give up some detail in describing data because they use only 5 summary numbers for comparing data sets. Note: The definition of lower and upper quartile used in this text is slightly different than the one used by some other authors (and software packages). Technically speaking, the median of the lower half of the data is not really the first quartile, although it is generally *very close*. Instead, the medians of the lower and upper halves of the data are often called the **lower** and **upper hinges**. Our boxplots use the lower and upper hinges to define the spread of the middle 50% of the data, but other authors sometimes use the *actual* quartiles for this purpose. The difference is usually very slight, usually unnoticeable, but not always. For example in the data sets of this exercise, a comparative boxplot based on the actual quartiles (as computed by Minitab) is shown below. The graph shows substantially the same type of information as those described in (a) except the graphs based on quartiles are able to detect the slight differences in variation between the three data sets.



Chapter 1: Overview and Descriptive Statistics

76. The measures that are sensitive to outliers are: the mean and the midrange. The mean is sensitive because all values are used in computing it. The midrange is sensitive because it uses only the most extreme values in its computation.

The median, the trimmed mean, and the midhinge are not sensitive to outliers.

The median is the most resistant to outliers because it uses only the middle value (or values) in its computation.

The trimmed mean is somewhat resistant to outliers. The larger the trimming percentage, the more resistant the trimmed mean becomes.

The midhinge, which uses the quartiles, is reasonably resistant to outliers because both quartiles are resistant to outliers.

77.

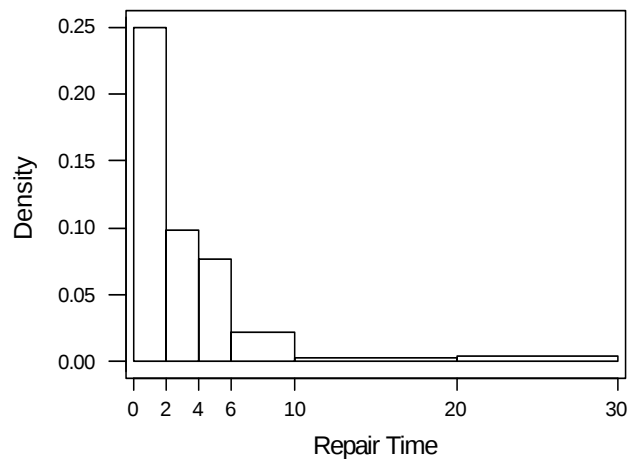
a.

0	2355566777888	
1	0000135555	
2	00257	
3	0033	
4	0057	
5	044	
6		stem: ones
7	05	leaf: tenths
8	8	
9	0	
10	3	
HI	22.0 24.5	

Chapter 1: Overview and Descriptive Statistics

b.

Interval	Frequency	Rel. Freq.	Density
0 -< 2	23	.500	.250
2 -< 4	9	.196	.098
4 -< 6	7	.152	.076
6 -< 10	4	.087	.022
10 -< 20	1	.022	.002
20 -< 30	2	.043	.004



78.

- Since the constant $\bar{x}$ is subtracted from each x value to obtain each y value, and addition or subtraction of a constant doesn't affect variability, $s_y^2 = s_x^2$ and $s_y = s_x$.
- Let $c = 1/s$, where s is the sample standard deviation of the x's and also (by a) of the y's. Then $s_z = cs_y = (1/s)s = 1$, and $s_z^2 = 1$. That is, the "standardized" quantities $z_1, \dots, z_n$ have a sample variance and standard deviation of 1.

79.

$$\text{a. } \sum_{i=1}^{n+1} x_i = \sum_{i=1}^n x_i + x_{n+1} = n\bar{x}_n + x_{n+1}, \text{ so } \bar{x}_{n+1} = \frac{[n\bar{x}_n + x_{n+1}]}{(n+1)}$$

b.

$$\begin{aligned} ns_{n+1}^2 &= \sum_{i=1}^{n+1} (x_i - \bar{x}_{n+1})^2 = \sum_{i=1}^{n+1} x_i^2 - (n+1)\bar{x}_{n+1}^2 \\ &= \sum_{i=1}^n x_i^2 - n\bar{x}_n^2 + x_{n+1}^2 + n\bar{x}_n^2 - (n+1)\bar{x}_{n+1}^2 \\ &= (n-1)s_n^2 + \{x_{n+1}^2 + n\bar{x}_n^2 - (n+1)\bar{x}_{n+1}^2\} \end{aligned}$$

When the expression for $\bar{x}_{n+1}$ from **a** is substituted, the expression in braces simplifies to

$$\text{the following, as desired: } \frac{n(x_{n+1} - \bar{x}_n)^2}{(n+1)}$$

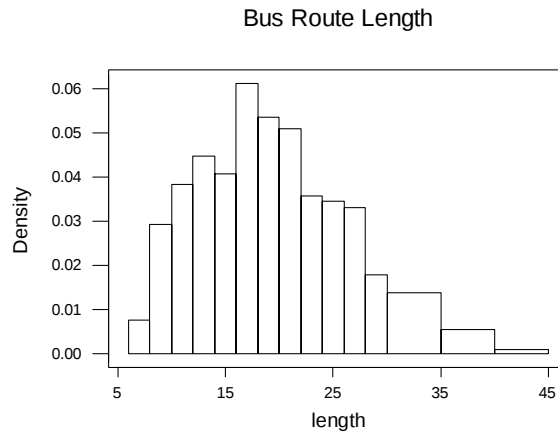
$$\text{c. } \bar{x}_{n+1} = \frac{15(12.58) + 11.8}{16} = \frac{200.5}{16} = 12.53$$

$$\begin{aligned} s_{n+1}^2 &= \frac{n-1}{n}(s_n^2) + \frac{(x_{n+1} - \bar{x}_n)^2}{(n+1)} = \frac{14}{15}(.512^2) + \frac{(11.8 - 12.58)^2}{(16)} \\ &= .245 + .038 = .283. \text{ So the standard deviation } s_{n+1} = \sqrt{.283} = .532 \end{aligned}$$

Chapter 1: Overview and Descriptive Statistics

80.

a.



b. Proportion less than 20 = $\left(\frac{216}{391}\right) = .552$

Proportion at least 30 = $\left(\frac{40}{391}\right) = .102$

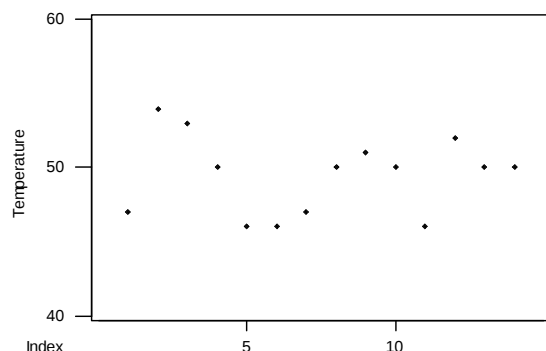
- c. First compute $(.90)(391 + 1) = 352.8$. Thus, the 90th percentile should be about the 352nd ordered value. The 351st ordered value lies in the interval 28 - < 30. The 352nd ordered value lies in the interval 30 - < 35. There are 27 values in the interval 30 - < 35. We do not know how these values are distributed, however, the smallest value (i.e., the 352nd value in the data set) cannot be smaller than 30. So, the 90th percentile is roughly 30.
- d. First compute $(.50)(391 + 1) = 196$. Thus the median (50th percentile) should be the 196th ordered value. The 174th ordered value lies in the interval 16 - < 18. The next 42 observations lie in the interval 18 - < 20. So, ordered observation 175 to 216 lie in the intervals 18 - < 20. The 196th observation is about in the middle of these. Thus, we would say, the median is roughly 19.

81.

Assuming that the histogram is unimodal, then there is evidence of positive skewness in the data since the median lies to the left of the mean (for a symmetric distribution, the mean and median would coincide). For more evidence of skewness, compare the distances of the 5th and 95th percentiles from the median: median - 5th percentile = 500 - 400 = 100 while 95th percentile - median = 720 - 500 = 220. Thus, the largest 5% of the values (above the 95th percentile) are further from the median than are the lowest 5%. The same skewness is evident when comparing the 10th and 90th percentiles to the median: median - 10th percentile = 500 - 430 = 70 while 90th percentile - median = 640 - 500 = 140. Finally, note that the largest value (925) is much further from the median (925 - 500 = 425) than is the smallest value (500 - 220 = 280), again an indication of positive skewness.

82.

- a. There is some evidence of a cyclical pattern.



- b.
- $$\bar{x}_2 = .1x_2 + .9\bar{x}_1 = (.1)(54) + (.9)(47) = 47.7$$
- $$\bar{x}_3 = .1x_3 + .9\bar{x}_2 = (.1)(53) + (.9)(47.7) = 48.23 \approx 48.2, \text{ etc.}$$

t	$\bar{x}_t$ for $\alpha = .1$	$\bar{x}_t$ for $\alpha = .5$
1	47.0	47.0
2	47.7	50.5
3	48.2	51.8
4	48.4	50.9
5	48.2	48.4
6	48.0	47.2
7	47.9	47.1
8	48.1	48.6
9	48.4	49.8
10	48.5	49.9
11	48.3	47.9
12	48.6	50.0
13	48.8	50.0
14	48.9	50.0

$\alpha = .1$ gives a smoother series.

- c.
- $$\begin{aligned} \bar{x}_t &= \alpha x_t + (1-\alpha)\bar{x}_{t-1} \\ &= \alpha x_t + (1-\alpha)[\alpha x_{t-1} + (1-\alpha)\bar{x}_{t-2}] \\ &= \alpha x_t + \alpha(1-\alpha)x_{t-1} + (1-\alpha)^2[\alpha x_{t-2} + (1-\alpha)\bar{x}_{t-3}] \\ &= \dots = \alpha x_t + \alpha(1-\alpha)x_{t-1} + \alpha(1-\alpha)^2x_{t-2} + \dots + \alpha(1-\alpha)^{t-2}x_2 + (1-\alpha)^{t-1}\bar{x}_1 \end{aligned}$$

Thus, $(\bar{x})_t$ depends on x_t and all previous values. As k increases, the coefficient on x_k decreases (further back in time implies less weight).

- d. Not very sensitive, since $(1-\alpha)^{t-1}$ will be very small.

Chapter 1: Overview and Descriptive Statistics

83.

- a. When there is perfect symmetry, the smallest observation y_1 and the largest observation y_n will be equidistant from the median, so $y_n - \bar{x} = \bar{x} - y_1$.
Similarly, the second smallest and second largest will be equidistant from the median, so $y_{n-1} - \bar{x} = \bar{x} - y_2$
and so on. Thus, the first and second numbers in each pair will be equal, so that each point in the plot will fall exactly on the 45 degree line. When the data is positively skewed, y_n will be much further from the median than is y_1 , so $y_n - \bar{x}$ will considerably exceed $\bar{x} - y_1$ and the point $(y_n - \bar{x}, \bar{x} - y_1)$ will fall considerably below the 45 degree line. A similar comment applies to other points in the plot.
- b. The first point in the plot is $(2745.6 - 221.6, 221.6 - 4.1) = (2524.0, 217.5)$. The others are: $(1476.2, 213.9)$, $(1434.4, 204.1)$, $(756.4, 190.2)$, $(481.8, 188.9)$, $(267.5, 181.0)$, $(208.4, 129.2)$, $(112.5, 106.3)$, $(81.2, 103.3)$, $(53.1, 102.6)$, $(53.1, 92.0)$, $(33.4, 23.0)$, and $(20.9, 20.9)$. The first number in each of the first seven pairs greatly exceed the second number, so each point falls well below the 45 degree line. A substantial positive skew (stretched upper tail) is indicated.

CHAPTER 2

Section 2.1

1.

- a. $S = \{ 1324, 1342, 1423, 1432, 2314, 2341, 2413, 2431, 3124, 3142, 4123, 4132, 3214, 3241, 4213, 4231 \}$
- b. Event A contains the outcomes where 1 is first in the list:
 $A = \{ 1324, 1342, 1423, 1432 \}$
- c. Event B contains the outcomes where 2 is first or second:
 $B = \{ 2314, 2341, 2413, 2431, 3214, 3241, 4213, 4231 \}$
- d. The compound event $A \cup B$ contains the outcomes in A or B or both:
 $A \cup B = \{ 1324, 1342, 1423, 1432, 2314, 2341, 2413, 2431, 3214, 3241, 4213, 4231 \}$

2.

- a. Event $A = \{ RRR, LLL, SSS \}$
- b. Event $B = \{ RLS, RSL, LRS, LSR, SRL, SLR \}$
- c. Event $C = \{ RRL, RRS, RLR, RSR, LRR, SRR \}$
- d. Event $D = \{ RRL, RRS, RLR, RSR, LRR, SRR, LLR, LLS, LRL, LSL, RLL, SLL, SSR, SSL, SRS, SLS, RSS, LSS \}$
- e. Event D' contains outcomes where all cars go the same direction, or they all go different directions:
 $D' = \{ RRR, LLL, SSS, RLS, RSL, LRS, LSR, SRL, SLR \}$

Because Event D totally encloses Event C, the compound event $C \cup D = D$:

$C \cup D = \{ RRL, RRS, RLR, RSR, LRR, SRR, LLR, LLS, LRL, LSL, RLL, SLL, SSR, SSL, SRS, SLS, RSS, LSS \}$

Using similar reasoning, we see that the compound event $C \cap D = C$:

$C \cap D = \{ RRL, RRS, RLR, RSR, LRR, SRR \}$

Chapter 2: Probability

3.

- a. Event $A = \{ SSF, SFS, FSS \}$
- b. Event $B = \{ SSS, SSF, SFS, FSS \}$
- c. For Event C, the system must have component 1 working (S in the first position), then at least one of the other two components must work (at least one S in the 2nd and 3rd positions: Event $C = \{ SSS, SSF, SFS \}$
- d. Event $C' = \{ SFF, FSS, FSF, FFS, FFF \}$
 Event $A \cup C = \{ SSS, SSF, SFS, FSS \}$
 Event $A \cap C = \{ SSF, SFS \}$
 Event $B \cup C = \{ SSS, SSF, SFS, FSS \}$
 Event $B \cap C = \{ SSS, SSF, SFS \}$

4.

a.

Outcome	Home Mortgage Number			
	1	2	3	4
1	F	F	F	F
2	F	F	F	V
3	F	F	V	F
4	F	F	V	V
5	F	V	F	F
6	F	V	F	V
7	F	V	V	F
8	F	V	V	V
9	V	F	F	F
10	V	F	F	V
11	V	F	V	F
12	V	F	V	V
13	V	V	F	F
14	V	V	F	V
15	V	V	V	F
16	V	V	V	V

- b. Outcome numbers 2, 3, 5, 9
- c. Outcome numbers 1, 16
- d. Outcome numbers 1, 2, 3, 5, 9
- e. In words, the UNION described is the event that either all of the mortgages are variable, or that at most all of them are variable: outcomes 1,2,3,5,9,16. The INTERSECTION described is the event that all of the mortgages are fixed: outcome 1.
- f. The UNION described is the event that either exactly three are fixed, or that all four are the same: outcomes 1, 2, 3, 5, 9, 16. The INTERSECTION in words is the event that exactly three are fixed AND that all four are the same. This cannot happen. (There are no outcomes in common) : $\mathbf{b} \cap \mathbf{c} = \emptyset$.

5.

a.

Outcome Number	Outcome
1	111
2	112
3	113
4	121
5	122
6	123
7	131
8	132
9	133
10	211
11	212
12	213
13	221
14	222
15	223
16	231
17	232
18	233
19	311
20	312
21	313
22	321
23	322
24	323
25	331
26	332
27	333

b. Outcome Numbers 1, 14, 27

c. Outcome Numbers 6, 8, 12, 16, 20, 22

d. Outcome Numbers 1, 3, 7, 9, 19, 21, 25, 27

Chapter 2: Probability

6.

a.

Outcome Number	Outcome
1	123
2	124
3	125
4	213
5	214
6	215
7	13
8	14
9	15
10	23
11	24
12	25
13	3
14	4
15	5

b. Outcomes 13, 14, 15

c. Outcomes 3, 6, 9, 12, 15

d. Outcomes 10, 11, 12, 13, 14, 15

7.

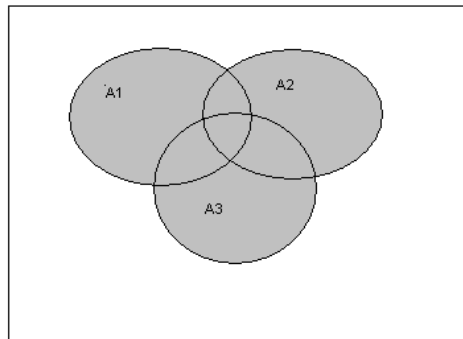
a. $S = \{BBBAAAA, BBABAAA, BBAABAA, BBAAABA, BBAAAAB, BABBAAA, BABABAA, BABAABA, BABAAAB, BAABBAA, BAABABA, BAABAAB, BAAABBA, BAAABAB, BAAAABB, ABBBAAA, ABBABAA, ABBAABA, ABBAaab, ABABBAA, ABABABA, ABABAAB, ABAABBA, ABAABAB, ABAAABB, AABBBAA, AABBABA, AABBAAB, AABABBA, AABABAB, AABAABB, AAABBBB, AAABBAB, AAABABB, AAAABBB\}$

b. $\{AAAABBB, AAABABB, AAABBAB, AABAABB, AABABAB\}$

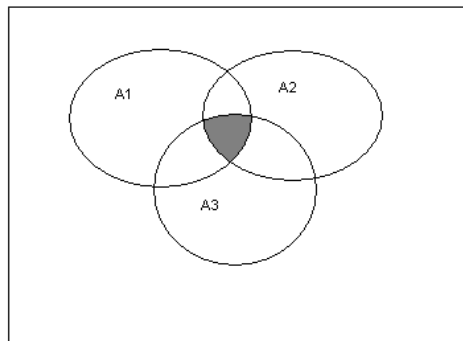
Chapter 2: Probability

8.

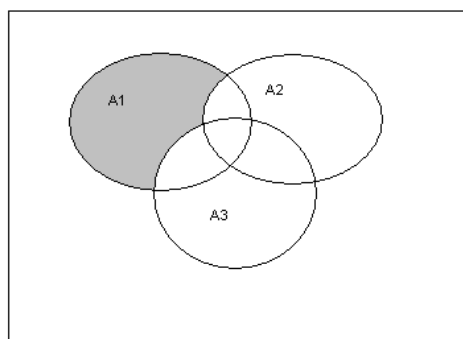
a. $A_1 \cup A_2 \cup A_3$



b. $A_1 \cap A_2 \cap A_3$

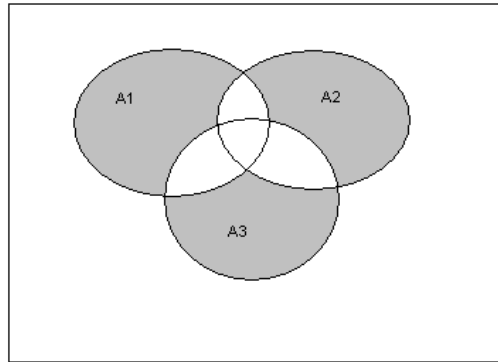


c. $A_1 \cap A_2' \cap A_3'$

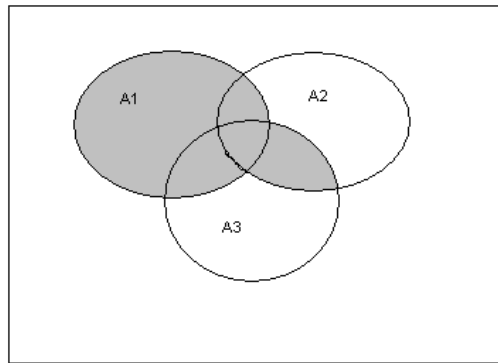


Chapter 2: Probability

d. $(A_1 \cap A_2' \cap A_3') \cup (A_1' \cap A_2 \cap A_3') \cup (A_1' \cap A_2' \cap A_3)$



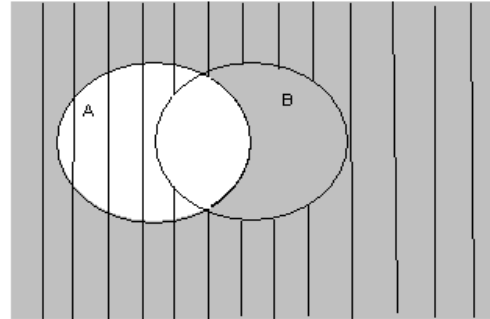
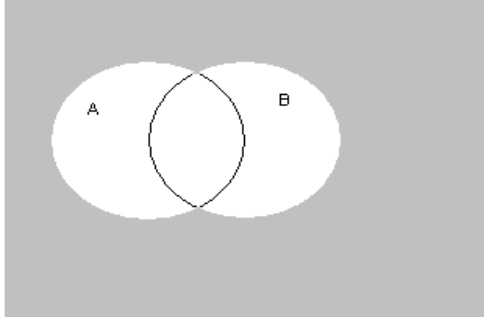
e. $A_1 \cup (A_2 \cap A_3)$



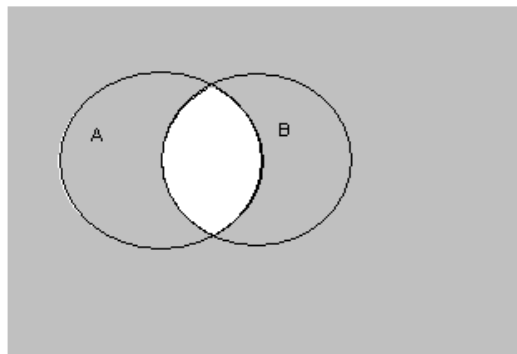
Chapter 2: Probability

9.

- a. In the diagram on the left, the shaded area is $(A \cup B)'$. On the right, the shaded area is A' , the striped area is B' , and the intersection $A' \cap B'$ occurs where there is BOTH shading and stripes. These two diagrams display the same area.



- b. In the diagram below, the shaded area represents $(A \cap B)'$. Using the diagram on the right above, the union of A' and B' is represented by the areas that have either shading or stripes. Both of the diagrams display the same area.



10.

- a. $A = \{\text{Chev, Pont, Buick}\}$, $B = \{\text{Ford, Merc}\}$, $C = \{\text{Plym, Chrys}\}$ are three mutually exclusive events.
- b. No, let $E = \{\text{Chev, Pont}\}$, $F = \{\text{Pont, Buick}\}$, $G = \{\text{Buick, Ford}\}$. These events are not mutually exclusive (e.g. E and F have an outcome in common), yet there is no outcome common to all three events.

Section 2.2

11.

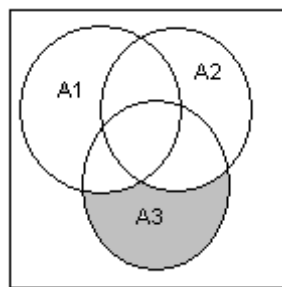
- a. .07
- b. $.15 + .10 + .05 = .30$
- c. Let event A = selected customer owns stocks. Then the probability that a selected customer does not own a stock can be represented by
 $P(A') = 1 - P(A) = 1 - (.18 + .25) = 1 - .43 = .57$. This could also have been done easily by adding the probabilities of the funds that are not stocks.

12.

- a. $P(A \cup B) = .50 + .40 - .25 = .65$
- b. $P(A \cup B)' = 1 - .65 = .35$
- c. $A \cap B'$; $P(A \cap B') = P(A) - P(A \cap B) = .50 - .25 = .25$

13.

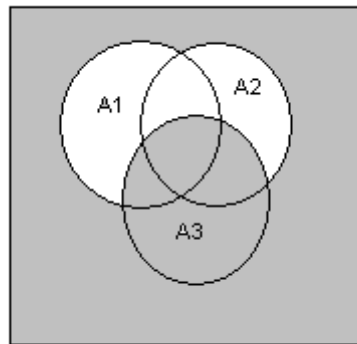
- a. awarded either #1 or #2 (or both):
 $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2) = .22 + .25 - .11 = .36$
- b. awarded neither #1 or #2:
 $P(A_1' \cap A_2') = P[(A_1 \cup A_2)'] = 1 - P(A_1 \cup A_2) = 1 - .36 = .64$
- c. awarded at least one of #1, #2, #3:
 $P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) - P(A_1 \cap A_2) - P(A_1 \cap A_3) - P(A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_3)$
 $= .22 + .25 + .28 - .11 - .05 - .07 + .01 = .53$
- d. awarded none of the three projects:
 $P(A_1' \cap A_2' \cap A_3') = 1 - P(\text{awarded at least one}) = 1 - .53 = .47$.
- e. awarded #3 but neither #1 nor #2:
 $P(A_1' \cap A_2' \cap A_3) = P(A_3) - P(A_1 \cap A_3) - P(A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_3)$
 $= .28 - .05 - .07 + .01 = .17$



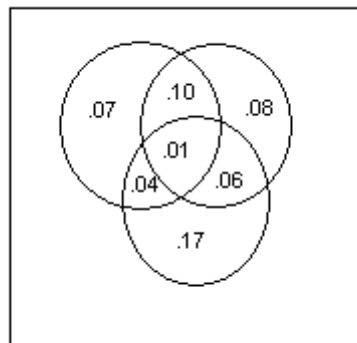
Chapter 2: Probability

- f. either (neither #1 nor #2) or #3:

$$\begin{aligned} P[(A_1' \cap A_2') \cup A_3] &= P(\text{shaded region}) = P(\text{awarded none}) + P(A_3) \\ &= .47 + .28 = .75 \end{aligned}$$



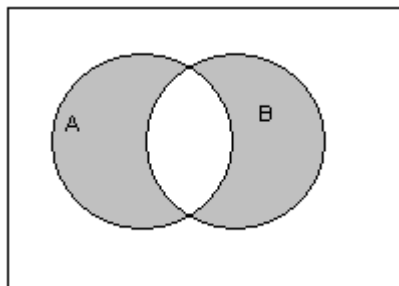
Alternatively, answers to **a – f** can be obtained from probabilities on the accompanying Venn diagram



Chapter 2: Probability

14.

- a. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$,
 so $P(A \cap B) = P(A) + P(B) - P(A \cup B)$
 $= .8 + .7 - .9 = .6$
- b. $P(\text{shaded region}) = P(A \cup B) - P(A \cap B) = .9 - .6 = .3$
 Shaded region = event of interest = $(A \cap B') \cup (A' \cap B)$



15.

- a. Let event E be the event that at most one purchases an electric dryer. Then E' is the event that at least two purchase electric dryers.
 $P(E') = 1 - P(E) = 1 - .428 = .572$
- b. Let event A be the event that all five purchase gas. Let event B be the event that all five purchase electric. All other possible outcomes are those in which at least one of each type is purchased. Thus, the desired probability =
 $1 - P(A) - P(B) = 1 - .116 - .005 = .879$

16.

- a. There are six simple events, corresponding to the outcomes CDP, CPD, DCP, DPC, PCD, and PDC. The probability assigned to each is $\frac{1}{6}$.
- b. $P(\text{C ranked first}) = P(\{CPD, CDP\}) = \frac{1}{6} + \frac{1}{6} = \frac{2}{6} = .333$
- c. $P(\text{C ranked first and D last}) = P(\{CPD\}) = \frac{1}{6}$

Chapter 2: Probability

17.

- a. The probabilities do not add to 1 because there are other software packages besides SPSS and SAS for which requests could be made.
- b. $P(A') = 1 - P(A) = 1 - .30 = .70$
- c. $P(A \cup B) = P(A) + P(B) = .30 + .50 = .80$
(since A and B are mutually exclusive events)
- d. $P(A' \cap B') = P[(A \cup B)']$ (De Morgan's law)
 $= 1 - P(A \cup B)$
 $= 1 - .80 = .20$

18.

This situation requires the complement concept. The only way for the desired event NOT to happen is if a 75 W bulb is selected first. Let event A be that a 75 W bulb is selected first, and $P(A) = \frac{6}{15}$. Then the desired event is event A' .

$$\text{So } P(A') = 1 - P(A) = 1 - \frac{6}{15} = \frac{9}{15} = .60$$

19.

Let event A be that the selected joint was found defective by inspector A. $P(A) = \frac{724}{10,000}$. Let event B be analogous for inspector B. $P(B) = \frac{751}{10,000}$. Compound event $A \cup B$ is the event that the selected joint was found defective by at least one of the two inspectors. $P(A \cup B) = \frac{1159}{10,000}$.

- a. The desired event is $(A \cup B)'$, so we use the complement rule:

$$P(A \cup B)' = 1 - P(A \cup B) = 1 - \frac{1159}{10,000} = \frac{8841}{10,000} = .8841$$

- b. The desired event is $B \cap A'$. $P(B \cap A') = P(B) - P(A \cap B)$.

$$P(A \cap B) = P(A) + P(B) - P(A \cup B),$$

$$= .0724 + .0751 - .1159 = .0316$$

$$\text{So } P(B \cap A') = P(B) - P(A \cap B)$$

$$= .0751 - .0316 = .0435$$

20.

Let S1, S2 and S3 represent the swing and night shifts, respectively. Let C1 and C2 represent the unsafe conditions and unrelated to conditions, respectively.

- a. The simple events are $\{S1, C1\}$, $\{S1, C2\}$, $\{S2, C1\}$, $\{S2, C2\}$, $\{S3, C1\}$, $\{S3, C2\}$.

- b. $P(\{C1\}) = P(\{S1, C1\}, \{S2, C1\}, \{S3, C1\}) = .10 + .08 + .05 = .23$

- c. $P(\{S1\}') = 1 - P(\{S1, C1\}, \{S1, C2\}) = 1 - (.10 + .35) = .55$

Chapter 2: Probability

21.

- a. $P(\{M,H\}) = .10$
- b. $P(\text{low auto}) = P[\{(L,N), (L,L), (L,M), (L,H)\}] = .04 + .06 + .05 + .03 = .18$ Following a similar pattern, $P(\text{low homeowner's}) = .06 + .10 + .03 = .19$
- c. $P(\text{same deductible for both}) = P[\{LL, MM, HH\}] = .06 + .20 + .15 = .41$
- d. $P(\text{deductibles are different}) = 1 - P(\text{same deductibles}) = 1 - .41 = .59$
- e. $P(\text{at least one low deductible}) = P[\{LN, LL, LM, LH, ML, HL\}]$
 $= .04 + .06 + .05 + .03 + .10 + .03 = .31$
- f. $P(\text{neither low}) = 1 - P(\text{at least one low}) = 1 - .31 = .69$

22.

- a. $P(A_1 \cap A_2) = P(A_1) + P(A_2) - P(A_1 \cup A_2) = .4 + .5 - .6 = .3$
- b. $P(A_1 \cap A_2') = P(A_1) - P(A_1 \cap A_2) = .4 - .3 = .1$
- c. $P(\text{exactly one}) = P(A_1 \cup A_2) - P(A_1 \cap A_2) = .6 - .3 = .3$

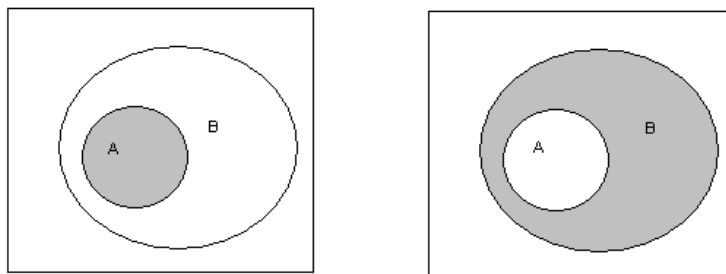
23.

Assume that the computers are numbered 1 – 6 as described. Also assume that computers 1 and 2 are the laptops. Possible outcomes are (1,2) (1,3) (1,4) (1,5) (1,6) (2,3) (2,4) (2,5) (2,6) (3,4) (3,5) (3,6) (4,5) (4,6) and (5,6).

- a. $P(\text{both are laptops}) = P[\{(1,2)\}] = \frac{1}{15} = .067$
- b. $P(\text{both are desktops}) = P[\{(3,4) (3,5) (3,6) (4,5) (4,6) (5,6)\}] = \frac{6}{15} = .40$
- c. $P(\text{at least one desktop}) = 1 - P(\text{no desktops})$
 $= 1 - P(\text{both are laptops})$
 $= 1 - .067 = .933$
- d. $P(\text{at least one of each type}) = 1 - P(\text{both are the same})$
 $= 1 - P(\text{both laptops}) - P(\text{both desktops})$
 $= 1 - .067 - .40 = .533$

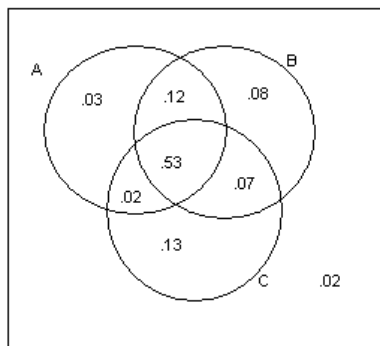
Chapter 2: Probability

- 24.** Since A is contained in B, then B can be written as the union of A and $(B \cap A')$, two mutually exclusive events. (See diagram).



From Axiom 3, $P[A \cup (B \cap A')] = P(A) + P(B \cap A')$. Substituting $P(B)$, $P(B) = P(A) + P(B \cap A')$ or $P(B) - P(A) = P(B \cap A')$. From Axiom 1, $P(B \cap A') \geq 0$, so $P(B) \geq P(A)$ or $P(A) \leq P(B)$. For general events A and B, $P(A \cap B) \leq P(A)$, and $P(A \cup B) \geq P(A)$.

- 25.** $P(A \cap B) = P(A) + P(B) - P(A \cup B) = .65$
 $P(A \cap C) = .55$, $P(B \cap C) = .60$
 $P(A \cap B \cap C) = P(A \cup B \cup C) - P(A) - P(B) - P(C)$
 $\quad + P(A \cap B) + P(A \cap C) + P(B \cap C)$
 $= .98 - .7 - .8 - .75 + .65 + .55 + .60$
 $= .53$



- a. $P(A \cup B \cup C) = .98$, as given.
- b. $P(\text{none selected}) = 1 - P(A \cup B \cup C) = 1 - .98 = .02$
- c. $P(\text{only automatic transmission selected}) = .03$ from the Venn Diagram
- d. $P(\text{exactly one of the three}) = .03 + .08 + .13 = .24$

Chapter 2: Probability

26.

- a. $P(A_1') = 1 - P(A_1) = 1 - .12 = .88$
- b. $P(A_1 \cap A_2) = P(A_1) + P(A_2) - P(A_1 \cup A_2) = .12 + .07 - .13 = .06$
- c. $P(A_1 \cap A_2 \cap A_3') = P(A_1 \cap A_2) - P(A_1 \cap A_2 \cap A_3) = .06 - .01 = .05$
- d. $P(\text{at most two errors}) = 1 - P(\text{all three types})$
 $= 1 - P(A_1 \cap A_2 \cap A_3)$
 $= 1 - .01 = .99$

27.

- Outcomes: (A,B) (A,C₁) (A,C₂) (A,F) (B,A) (B,C₁) (B,C₂) (B,F)
 (C₁,A) (C₁,B) (C₁,C₂) (C₁,F) (C₂,A) (C₂,B) (C₂,C₁) (C₂,F)
 (F,A) (F,B) (F,C₁) (F,C₂)
- a. $P[(A,B) \text{ or } (B,A)] = \frac{2}{20} = \frac{1}{10} = .1$
 - b. $P(\text{at least one C}) = \frac{14}{20} = \frac{7}{10} = .7$
 - c. $P(\text{at least 15 years}) = 1 - P(\text{at most 14 years})$
 $= 1 - P[(3,6) \text{ or } (6,3) \text{ or } (3,7) \text{ or } (7,3) \text{ or } (3,10) \text{ or } (10,3) \text{ or } (6,7) \text{ or } (7,6)]$
 $= 1 - \frac{8}{20} = 1 - .4 = .6$

28.

- There are 27 equally likely outcomes.
- a. $P(\text{all the same}) = P[(1,1,1) \text{ or } (2,2,2) \text{ or } (3,3,3)] = \frac{3}{27} = \frac{1}{9}$
 - b. $P(\text{at most 2 are assigned to the same station}) = 1 - P(\text{all 3 are the same})$
 $= 1 - \frac{3}{27} = \frac{24}{27} = \frac{8}{9}$
 - c. $P(\text{all different}) = [\{ (1,2,3) (1,3,2) (2,1,3) (2,3,1) (3,1,2) (3,2,1) \}]$
 $= \frac{6}{27} = \frac{2}{9}$

Chapter 2: Probability

Section 2.3

29.

- a. $(5)(4) = 20$ (5 choices for president, 4 remain for vice president)
- b. $(5)(4)(3) = 60$
- c. $\binom{5}{2} = \frac{5!}{2!3!} = 10$ (No ordering is implied in the choice)

30.

- a. Because order is important, we'll use $P_{8,3} = 8(7)(6) = 336$.
- b. Order doesn't matter here, so we use $C_{30,6} = 593,775$.
- c. From each group we choose 2: $\binom{8}{2} \cdot \binom{10}{2} \cdot \binom{12}{2} = 83,160$
- d. The numerator comes from part c and the denominator from part b: $\frac{83,160}{593,775} = .14$
- e. We use the same denominator as in part d. We can have all zinfandel, all merlot, or all cabernet, so $P(\text{all same}) = P(\text{all z}) + P(\text{all m}) + P(\text{all c}) =$
$$\frac{\binom{8}{6} + \binom{10}{6} + \binom{12}{6}}{\binom{30}{6}} = \frac{1162}{593,775} = .002$$

31.

- a. $(n_1)(n_2) = (9)(27) = 243$
- b. $(n_1)(n_2)(n_3) = (9)(27)(15) = 3645$, so such a policy could be carried out for 3645 successive nights, or approximately 10 years, without repeating exactly the same program.

Chapter 2: Probability

32.

a. $5 \times 4 \times 3 \times 4 = 240$

b. $1 \times 1 \times 3 \times 4 = 12$

c. $4 \times 3 \times 3 \times 3 = 108$

d. # with at least on Sony = total # - # with no Sony = $240 - 108 = 132$

e. $P(\text{at least one Sony}) = \frac{132}{240} = .55$

$$\begin{aligned} P(\text{exactly one Sony}) &= P(\text{only Sony is receiver}) \\ &\quad + P(\text{only Sony is CD player}) \\ &\quad + P(\text{only Sony is deck}) \\ &= \frac{1 \times 3 \times 3 \times 3}{240} + \frac{4 \times 1 \times 3 \times 3}{240} + \frac{4 \times 3 \times 3 \times 1}{240} = \frac{27 + 36 + 36}{240} \\ &= \frac{99}{240} = .413 \end{aligned}$$

33.

a. $\binom{25}{5} = \frac{25!}{5!20!} = 53,130$

b. $\binom{8}{4} \cdot \binom{17}{1} = 1190$

c. $P(\text{exactly 4 have cracks}) = \frac{\binom{8}{4} \binom{17}{1}}{\binom{25}{5}} = \frac{1190}{53,130} = .022$

d. $P(\text{at least 4}) = P(\text{exactly 4}) + P(\text{exactly 5})$

$$= \frac{\binom{8}{4} \binom{17}{1}}{\binom{25}{5}} + \frac{\binom{8}{5} \binom{17}{0}}{\binom{25}{5}} = .022 + .001 = .023$$

34.

$$\text{a. } \binom{20}{6} = 38,760. \quad P(\text{all from day shift}) = \frac{\binom{20}{6} \binom{25}{0}}{\binom{45}{6}} = \frac{38,760}{8,145,060} = .0048$$

$$\begin{aligned} \text{b. } P(\text{all from same shift}) &= \frac{\binom{20}{6} \binom{25}{0}}{\binom{45}{6}} + \frac{\binom{15}{6} \binom{30}{0}}{\binom{45}{6}} + \frac{\binom{10}{6} \binom{35}{0}}{\binom{45}{6}} \\ &= .0048 + .0006 + .0000 = .0054 \end{aligned}$$

$$\begin{aligned} \text{c. } P(\text{at least two shifts represented}) &= 1 - P(\text{all from same shift}) \\ &= 1 - .0054 = .9946 \end{aligned}$$

d. Let A_1 = day shift unrepresented, A_2 = swing shift unrepresented, and A_3 = graveyard shift unrepresented. Then we wish $P(A_1 \cup A_2 \cup A_3)$.

$P(A_1)$ = P(day unrepresented) = P(all from swing and graveyard)

$$P(A_1) = \frac{\binom{25}{6}}{\binom{45}{6}}, \quad P(A_2) = \frac{\binom{30}{6}}{\binom{45}{6}}, \quad P(A_3) = \frac{\binom{35}{6}}{\binom{45}{6}},$$

$$P(A_1 \cap A_2) = P(\text{all from graveyard}) = \frac{\binom{10}{6}}{\binom{45}{6}}$$

$$P(A_1 \cap A_3) = \frac{\binom{15}{6}}{\binom{45}{6}}, \quad P(A_2 \cap A_3) = \frac{\binom{20}{6}}{\binom{45}{6}}, \quad P(A_1 \cap A_2 \cap A_3) = 0,$$

$$\begin{aligned} \text{So } P(A_1 \cup A_2 \cup A_3) &= \frac{\binom{25}{6}}{\binom{45}{6}} + \frac{\binom{30}{6}}{\binom{45}{6}} + \frac{\binom{35}{6}}{\binom{45}{6}} - \frac{\binom{10}{6}}{\binom{45}{6}} - \frac{\binom{15}{6}}{\binom{45}{6}} - \frac{\binom{20}{6}}{\binom{45}{6}} \\ &= .2939 - .0054 = .2885 \end{aligned}$$

Chapter 2: Probability

35. There are 10 possible outcomes -- $\binom{5}{2}$ ways to select the positions for B's votes: BBAAA, BABAA, BAABA, BAAAB, ABBA, ABABA, ABAAB, AABBA, AABAB, and AAABB. Only the last two have A ahead of B throughout the vote count. Since the outcomes are equally likely, the desired probability is $\frac{2}{10} = .20$.

36.

- a. $n_1 = 3, n_2 = 4, n_3 = 5$, so $n_1 \times n_2 \times n_3 = 60$ runs
- b. $n_1 = 1$, (just one temperature), $n_2 = 2, n_3 = 5$ implies that there are 10 such runs.

37. There are $\binom{60}{5}$ ways to select the 5 runs. Each catalyst is used in 12 different runs, so the number of ways of selecting one run from each of these 5 groups is 12^5 . Thus the desired probability is $\frac{12^5}{\binom{60}{5}} = .0456$.

38.

a.
$$P(\text{selecting 2 - 75 watt bulbs}) = \frac{\binom{6}{2}\binom{9}{1}}{\binom{15}{3}} = \frac{15 \cdot 9}{455} = .2967$$

b.
$$P(\text{all three are the same}) = \frac{\binom{4}{3} + \binom{5}{3} + \binom{6}{3}}{\binom{15}{3}} = \frac{4 + 10 + 20}{455} = .0747$$

c.
$$\frac{\binom{4}{1}\binom{5}{1}\binom{6}{1}}{\binom{15}{3}} = \frac{120}{455} = .2637$$

Chapter 2: Probability

- d. To examine exactly one, a 75 watt bulb must be chosen first. (6 ways to accomplish this). To examine exactly two, we must choose another wattage first, then a 75 watt. (9×6 ways). Following the pattern, for exactly three, $9 \times 8 \times 6$ ways; for four, $9 \times 8 \times 7 \times 6$; for five, $9 \times 8 \times 7 \times 6 \times 6$.

$$\begin{aligned} P(\text{examine at least 6 bulbs}) &= 1 - P(\text{examine 5 or less}) \\ &= 1 - P(\text{examine exactly 1 or 2 or 3 or 4 or 5}) \\ &= 1 - [P(\text{one}) + P(\text{two}) + \dots + P(\text{five})] \end{aligned}$$

$$= 1 - \left[\frac{6}{15} + \frac{9 \times 6}{15 \times 14} + \frac{9 \times 8 \times 6}{15 \times 14 \times 13} + \frac{9 \times 8 \times 7 \times 6}{15 \times 14 \times 13 \times 12} + \frac{9 \times 8 \times 7 \times 6 \times 6}{15 \times 14 \times 13 \times 12 \times 11} \right]$$

$$\begin{aligned} &= 1 - [.4 + .2571 + .1582 + .0923 + .0503] \\ &= 1 - .9579 = .0421 \end{aligned}$$

39.

- a. We want to choose all of the 5 cordless, and 5 of the 10 others, to be among the first 10

serviced, so the desired probability is $\frac{\binom{5}{5} \binom{10}{5}}{\binom{15}{10}} = \frac{252}{3003} = .0839$

- b. Isolating one group, say the cordless phones, we want the other two groups represented in the last 5 serviced. So we choose 5 of the 10 others, except that we don't want to include the outcomes where the last five are all the same.

So we have $\frac{\binom{10}{5} - 2}{\binom{15}{5}}$. But we have three groups of phones, so the desired probability is

$$\frac{3 \cdot \left[\binom{10}{5} - 2 \right]}{\binom{15}{5}} = \frac{3(250)}{3003} = .2498.$$

- c. We want to choose 2 of the 5 cordless, 2 of the 5 cellular, and 2 of the corded phones:

$$\frac{\binom{5}{2} \binom{5}{2} \binom{5}{2}}{\binom{15}{6}} = \frac{1000}{5005} = .1998$$

Chapter 2: Probability

40.

- a. If the A's are distinguishable from one another, and similarly for the B's, C's and D's, then there are $12!$ Possible chain molecules. Six of these are:
 $A_1A_2A_3B_2C_3C_1D_3C_2D_1D_2B_3B_1$, $A_1A_3A_2B_2C_3C_1D_3C_2D_1D_2B_3B_1$
 $A_2A_1A_3B_2C_3C_1D_3C_2D_1D_2B_3B_1$, $A_2A_3A_1B_2C_3C_1D_3C_2D_1D_2B_3B_1$
 $A_3A_1A_2B_2C_3C_1D_3C_2D_1D_2B_3B_1$, $A_3A_2A_1B_2C_3C_1D_3C_2D_1D_2B_3B_1$
 These 6 ($=3!$) differ only with respect to ordering of the 3 A's. In general, groups of 6 chain molecules can be created such that within each group only the ordering of the A's is different. When the A subscripts are suppressed, each group of 6 "collapses" into a single molecule (B's, C's and D's are still distinguishable). At this point there are $\frac{12!}{3!}$ molecules. Now suppressing subscripts on the B's, C's and D's in turn gives ultimately $\frac{12!}{(3!)^4} = 369,600$ chain molecules.
- b. Think of the group of 3 A's as a single entity, and similarly for the B's, C's, and D's. Then there are $4!$ Ways to order these entities, and thus $4!$ Molecules in which the A's are contiguous, the B's, C's, and D's are also. Thus, $P(\text{all together}) = \frac{4!}{369,600} = .00006494$.

41.

- a. $P(\text{at least one F among } 1^{\text{st}} 3) = 1 - P(\text{no F's among } 1^{\text{st}} 3)$

$$= 1 - \frac{4 \times 3 \times 2}{8 \times 7 \times 6} = 1 - \frac{24}{336} = 1 - .0714 = .9286$$

An alternative method to calculate $P(\text{no F's among } 1^{\text{st}} 3)$ would be to choose none of the females and 3 of the 4 males, as follows:

$$\frac{\binom{4}{0} \binom{4}{3}}{\binom{8}{3}} = \frac{4}{56} = .0714, \text{ obviously producing the same result.}$$

- b. $P(\text{all F's among } 1^{\text{st}} 5) = \frac{\binom{4}{4} \binom{4}{1}}{\binom{8}{5}} = \frac{4}{56} = .0714$

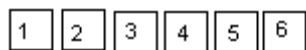
- c. $P(\text{orderings are different}) = 1 - P(\text{orderings are the same for both semesters})$

$$= 1 - \frac{(\# \text{ orderings such that the orders are the same each semester})}{(\text{total \# of possible orderings for 2 semesters})}$$

$$= 1 - \frac{8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1}{(8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1) \times (8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1)} = .99997520$$

Chapter 2: Probability

42. Seats:



$$P(\text{J\&P in 1\&2}) = \frac{2 \times 1 \times 4 \times 3 \times 2 \times 1}{6 \times 5 \times 4 \times 3 \times 2 \times 1} = \frac{1}{15} = .0667$$

$$\begin{aligned} P(\text{J\&P next to each other}) &= P(\text{J\&P in 1\&2}) + \dots + P(\text{J\&P in 5\&6}) \\ &= 5 \times \frac{1}{15} = \frac{1}{3} = .333 \end{aligned}$$

$$P(\text{at least one H next to his W}) = 1 - P(\text{no H next to his W})$$

We count the # of ways of no H next to his W as follows:

if orderings without a H-W pair in seats #1 and 3 and no H next to his W = $6^* \times 4 \times 1^* \times 2^{\#} \times 1 \times 1 = 48$

*= pair, # = can't put the mate of seat #2 here or else a H-W pair would be in #5 and 6.

of orderings without a H-W pair in seats #1 and 3, and no H next to his W = $6 \times 4 \times 2^{\#} \times 2 \times 2 \times 1 = 192$

= can't be mate of person in seat #1 or #2.

So, # of seating arrangements with no H next to W = $48 + 192 = 240$

$$\text{And } P(\text{no H next to his W}) = \frac{240}{6 \times 5 \times 4 \times 3 \times 2 \times 1} = \frac{1}{3}, \text{ so}$$

$$P(\text{at least one H next to his W}) = 1 - \frac{1}{3} = \frac{2}{3}$$

43. # of 10 high straights = $4 \times 4 \times 4 \times 4 \times 4$ (4 – 10's, 4 – 9's, etc)

$$P(10 \text{ high straight}) = \frac{4^5}{\binom{52}{5}} = \frac{1024}{2,598,960} = .000394$$

$$P(\text{straight}) = 10 \times \frac{4^5}{\binom{52}{5}} = .003940 \text{ (Multiply by 10 because there are 10 different card$$

values that could be high: Ace, King, etc.) There are only 40 straight flushes (10 in each suit), so

$$P(\text{straight flush}) = \frac{40}{\binom{52}{5}} = .00001539$$

$$44. \binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n!}{(n-k)!k!} = \binom{n}{n-k}$$

The number of subsets of size k = the number of subsets of size $n-k$, because to each subset of size k there corresponds exactly one subset of size $n-k$ (the $n-k$ objects not in the subset of size k).

Section 2.4

45.

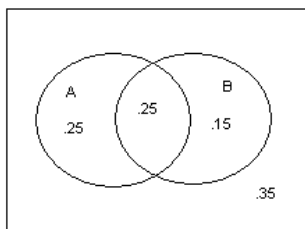
a. $P(A) = .106 + .141 + .200 = .447$, $P(C) = .215 + .200 + .065 + .020 = .500$ $P(A \cap C) = .200$

b. $P(A|C) = \frac{P(A \cap C)}{P(C)} = \frac{.200}{.500} = .400$. If we know that the individual came from ethnic group 3, the probability that he has type A blood is .40. $P(C|A) = \frac{P(A \cap C)}{P(A)} = \frac{.200}{.447} = .447$. If a person has type A blood, the probability that he is from ethnic group 3 is .447

c. Define event $D = \{\text{ethnic group 1 selected}\}$. We are asked for $P(D|B') = \frac{P(D \cap B')}{P(B')} = \frac{.200}{.500} = .400$. $P(D \cap B') = .082 + .106 + .004 = .192$, $P(B') = 1 - P(B) = 1 - [.008 + .018 + .065] = .909$

46. Let event A be that the individual is more than 6 feet tall. Let event B be that the individual is a professional basketball player. Then $P(A|B)$ = the probability of the individual being more than 6 feet tall, knowing that the individual is a professional basketball player, and $P(B|A)$ = the probability of the individual being a professional basketball player, knowing that the individual is more than 6 feet tall. $P(A|B)$ will be larger. Most professional BB players are tall, so the probability of an individual in that reduced sample space being more than 6 feet tall is very large. The number of individuals that are pro BB players is small in relation to the # of males more than 6 feet tall.

47.



- a. $P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{.25}{.50} = .50$
- b. $P(B'|A) = \frac{P(A \cap B')}{P(A)} = \frac{.25}{.50} = .50$
- c. $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{.25}{.40} = .6125$
- d. $P(A'|B) = \frac{P(A' \cap B)}{P(B)} = \frac{.15}{.40} = .3875$
- e. $P(A|A \cup B) = \frac{P[A \cap (A \cup B)]}{P(A \cup B)} = \frac{.50}{.65} = .7692$

48.

- a. $P(A_2|A_1) = \frac{P(A_1 \cap A_2)}{P(A_1)} = \frac{.06}{.12} = .50$
- b. $P(A_1 \cap A_2 \cap A_3|A_1) = \frac{.01}{.12} = .0833$

- c. We want $P[(\text{exactly one}) | (\text{at least one})]$.

$$\begin{aligned} P(\text{at least one}) &= P(A_1 \cup A_2 \cup A_3) \\ &= .12 + .07 + .05 - .06 - .03 - .02 + .01 = .14 \end{aligned}$$

Also notice that the intersection of the two events is just the 1st event, since “exactly one” is totally contained in “at least one.”

$$\text{So } P[(\text{exactly one}) | (\text{at least one})] = \frac{.04 + .01}{.14} = .3571$$

- d. The pieces of this equation can be found in your answers to exercise 26 (section 2.2):

$$P(A'_3 | A_1 \cap A_2) = \frac{P(A_1 \cap A_2 \cap A'_3)}{P(A_1 \cap A_2)} = \frac{.05}{.06} = .833$$

Chapter 2: Probability

- 49.** The first desired probability is $P(\text{both bulbs are 75 watt} \mid \text{at least one is 75 watt})$.
 $P(\text{at least one is 75 watt}) = 1 - P(\text{none are 75 watt})$

$$= 1 - \frac{\binom{9}{2}}{\binom{15}{2}} = 1 - \frac{36}{105} = \frac{69}{105}.$$

Notice that $P[(\text{both are 75 watt}) \cap (\text{at least one is 75 watt})]$

$$= P(\text{both are 75 watt}) = \frac{\binom{6}{2}}{\binom{15}{2}} = \frac{15}{105}.$$

$$\text{So } P(\text{both bulbs are 75 watt} \mid \text{at least one is 75 watt}) = \frac{\frac{15}{105}}{\frac{69}{105}} = \frac{15}{69} = .2174$$

Second, we want $P(\text{same rating} \mid \text{at least one NOT 75 watt})$.

$P(\text{at least one NOT 75 watt}) = 1 - P(\text{both are 75 watt})$

$$= 1 - \frac{15}{105} = \frac{90}{105}.$$

Now, $P[(\text{same rating}) \cap (\text{at least one not 75 watt})] = P(\text{both 40 watt or both 60 watt})$.

$$P(\text{both 40 watt or both 60 watt}) = \frac{\binom{4}{2} + \binom{5}{2}}{\binom{15}{2}} = \frac{16}{105}$$

$$\text{Now, the desired conditional probability is } \frac{\frac{16}{105}}{\frac{90}{105}} = \frac{16}{90} = .1778$$

50.

- a.** $P(M \cap LS \cap PR) = .05$, directly from the table of probabilities
- b.** $P(M \cap Pr) = P(M, Pr, LS) + P(M, Pr, SS) = .05 + .07 = .12$
- c.** $P(SS) = \text{sum of 9 probabilities in SS table} = .56$, $P(LS) = 1 - .56 = .44$
- d.** $P(M) = .08 + .07 + .12 + .10 + .05 + .07 = .49$
 $P(Pr) = .02 + .07 + .07 + .02 + .05 + .02 = .25$

Chapter 2: Probability

$$\text{e. } P(M|SS \cap PI) = \frac{P(M \cap SS \cap PI)}{P(SS \cap PI)} = \frac{.08}{.04 + .08 + .03} = .533$$

$$\text{f. } P(SS|M \cap PI) = \frac{P(SS \cap M \cap PI)}{P(M \cap PI)} = \frac{.08}{.08 + .10} = .444$$

$$P(LS|M \cap PI) = 1 - P(SS|M \cap PI) = 1 - .444 = .556$$

51.

$$\begin{aligned} \text{a. } P(R \text{ from } 1^{\text{st}} \cap R \text{ from } 2^{\text{nd}}) &= P(R \text{ from } 2^{\text{nd}} | R \text{ from } 1^{\text{st}}) \cdot P(R \text{ from } 1^{\text{st}}) \\ &= \frac{8}{11} \cdot \frac{6}{10} = .436 \end{aligned}$$

$$\begin{aligned} \text{b. } P(\text{same numbers}) &= P(\text{both selected balls are the same color}) \\ &= P(\text{both red}) + P(\text{both green}) = .436 + \frac{4}{11} \cdot \frac{4}{10} = .581 \end{aligned}$$

52. Let A_1 be the event that #1 fails and A_2 be the event that #2 fails. We assume that $P(A_1) = P(A_2) = q$ and that $P(A_1 | A_2) = P(A_2 | A_1) = r$. Then one approach is as follows:

$$P(A_1 \cap A_2) = P(A_2 | A_1) \cdot P(A_1) = rq = .01$$

$$P(A_1 \cup A_2) = P(A_1 \cap A_2) + P(A_1' \cap A_2) + P(A_1 \cap A_2') = rq + 2(1-r)q = .07$$

These two equations give $2q - .01 = .07$, from which $q = .04$ and $r = .25$. Alternatively, with $t = P(A_1' \cap A_2) = P(A_1 \cap A_2')$, $t + .01 + t = .07$, implying $t = .03$ and thus $q = .04$ without reference to conditional probability.

$$\begin{aligned} \text{53. } P(B|A) &= \frac{P(A \cap B)}{P(A)} = \frac{P(B)}{P(A)} \quad (\text{since } B \text{ is contained in } A, A \cap B = B) \\ &= \frac{.05}{.60} = .0833 \end{aligned}$$

Chapter 2: Probability

54. $P(A_1) = .22$, $P(A_2) = .25$, $P(A_3) = .28$, $P(A_1 \cap A_2) = .11$, $P(A_1 \cap A_3) = .05$, $P(A_2 \cap A_3) = .07$, $P(A_1 \cap A_2 \cap A_3) = .01$

a. $P(A_2 | A_1) = \frac{P(A_1 \cap A_2)}{P(A_1)} = \frac{.11}{.22} = .50$

b. $P(A_2 \cap A_3 | A_1) = \frac{P(A_1 \cap A_2 \cap A_3)}{P(A_1)} = \frac{.01}{.22} = .0455$

c.
$$P(A_2 \cup A_3 | A_1) = \frac{P[A_1 \cap (A_2 \cup A_3)]}{P(A_1)} = \frac{P[(A_1 \cap A_2) \cup (A_1 \cap A_3)]}{P(A_1)}$$

$$= \frac{P(A_1 \cap A_2) + P(A_1 \cap A_3) - P(A_1 \cap A_2 \cap A_3)}{P(A_1)} = \frac{.15}{.22} = .682$$

d. $P(A_1 \cap A_2 \cap A_3 | A_1 \cup A_2 \cup A_3) = \frac{P(A_1 \cap A_2 \cap A_3)}{P(A_1 \cup A_2 \cup A_3)} = \frac{.01}{.53} = .0189$

This is the probability of being awarded all three projects given that at least one project was awarded.

55.

a. $P(A \cap B) = P(B|A) \cdot P(A) = \frac{2 \times 1}{4 \times 3} \times \frac{2 \times 1}{6 \times 5} = .0111$

- b. $P(\text{two other H's next to their wives} | J \text{ and } M \text{ together in the middle})$

$$\frac{P[(H - W \text{ or } W - H) \text{ and } (J - M \text{ or } M - J) \text{ and } (H - W \text{ or } W - H)]}{P(J - M \text{ or } M - J \text{ in the middle})}$$

$$\text{numerator} = \frac{4 \times 1 \times 2 \times 1 \times 2 \times 1}{6 \times 5 \times 4 \times 3 \times 2 \times 1} = \frac{16}{6!}$$

$$\text{denominator} = \frac{4 \times 3 \times 2 \times 1 \times 2 \times 1}{6 \times 5 \times 4 \times 3 \times 2 \times 1} = \frac{48}{6!}$$

$$\text{so the desired probability} = \frac{16}{48} = \frac{1}{3}$$

Chapter 2: Probability

- c. $P(\text{all H's next to W's} \mid \text{J \& M together})$
 $= P(\text{all H's next to W's} - \text{including J\&M}) / P(\text{J\&M together})$

$$= \frac{\frac{6 \times 1 \times 4 \times 1 \times 2 \times 1}{6!}}{\frac{5 \times 2 \times 1 \times 4 \times 3 \times 2 \times 1}{6!}} = \frac{48}{240} = .2$$

56. If $P(B|A) > P(B)$, then $P(B'|A) < P(B')$.

Proof by contradiction.

Assume $P(B'|A) \geq P(B')$.

Then $1 - P(B|A) \geq 1 - P(B)$.

$$- P(B|A) \geq - P(B).$$

$$P(B|A) \leq P(B).$$

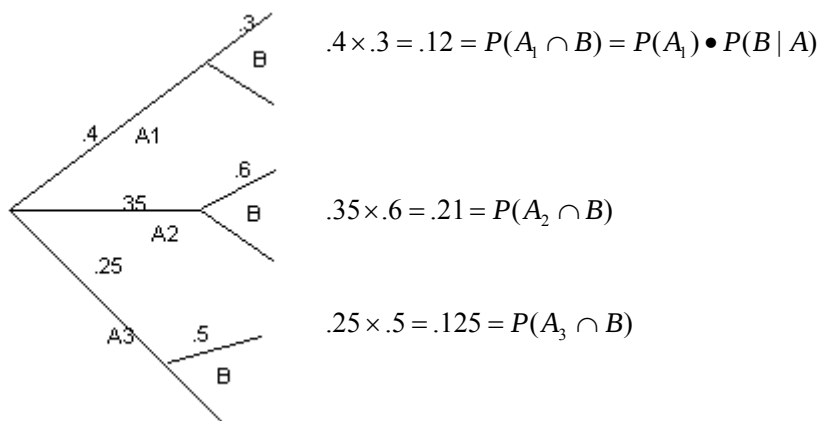
This contradicts the initial condition, therefore $P(B'|A) < P(B')$.

$$57. \quad P(A|B) + P(A'|B) = \frac{P(A \cap B)}{P(B)} + \frac{P(A' \cap B)}{P(B)} = \frac{P(A \cap B) + P(A' \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$$

$$\begin{aligned} 58. \quad P(A \cup B | C) &= \frac{P[(A \cup B) \cap C]}{P(C)} = \frac{P[(A \cap C) \cup (B \cap C)]}{P(C)} \\ &= \frac{P(A \cap C) + P(B \cap C) - P(A \cap B \cap C)}{P(C)} \\ &= P(A|C) + P(B|C) - P(A \cap B | C) \end{aligned}$$

Chapter 2: Probability

59.



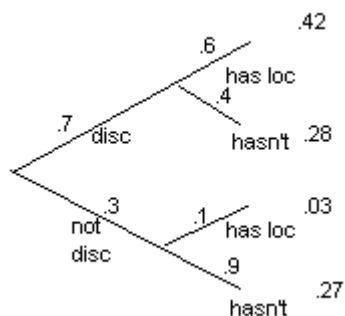
a. $P(A_2 \cap B) = .21$

b. $P(B) = P(A_1 \cap B) + P(A_2 \cap B) + P(A_3 \cap B) = .455$

c. $P(A_1|B) = \frac{P(A_1 \cap B)}{P(B)} = \frac{.12}{.455} = .264$

$P(A_2|B) = \frac{.21}{.455} = .462$, $P(A_3|B) = 1 - .264 - .462 = .274$

60.



a. $P(\text{not disc} | \text{has loc}) = \frac{P(\text{not disc} \cap \text{has loc})}{P(\text{has loc})} = \frac{.03}{.03 + .42} = .067$

b. $P(\text{disc} | \text{no loc}) = \frac{P(\text{disc} \cap \text{no loc})}{P(\text{no loc})} = \frac{.28}{.55} = .509$

Chapter 2: Probability

61. $P(0 \text{ def in sample} \mid 0 \text{ def in batch}) = 1$

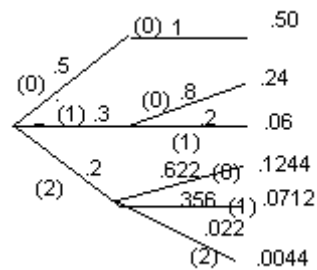
$$P(0 \text{ def in sample} \mid 1 \text{ def in batch}) = \frac{\binom{9}{2}}{\binom{10}{2}} = .800$$

$$P(1 \text{ def in sample} \mid 1 \text{ def in batch}) = \frac{\binom{9}{1}}{\binom{10}{2}} = .200$$

$$P(0 \text{ def in sample} \mid 2 \text{ def in batch}) = \frac{\binom{8}{2}}{\binom{10}{2}} = .622$$

$$P(1 \text{ def in sample} \mid 2 \text{ def in batch}) = \frac{\binom{2}{1}\binom{8}{1}}{\binom{10}{2}} = .356$$

$$P(2 \text{ def in sample} \mid 2 \text{ def in batch}) = \frac{1}{\binom{10}{2}} = .022$$



a. $P(0 \text{ def in batch} \mid 0 \text{ def in sample}) = \frac{.5}{.5 + .24 + .1244} = .578$

$$P(1 \text{ def in batch} \mid 0 \text{ def in sample}) = \frac{.24}{.5 + .24 + .1244} = .278$$

$$P(2 \text{ def in batch} \mid 0 \text{ def in sample}) = \frac{.1244}{.5 + .24 + .1244} = .144$$

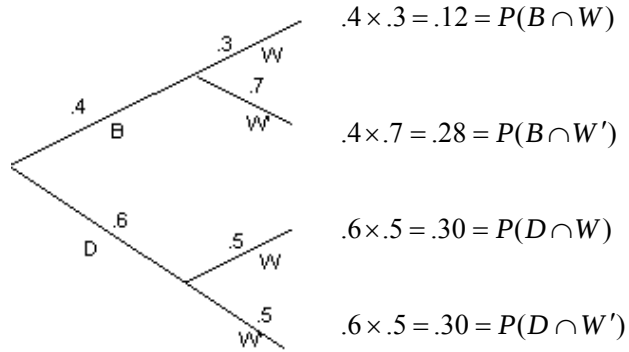
Chapter 2: Probability

b. $P(0 \text{ def in batch} \mid 1 \text{ def in sample}) = 0$

$$P(1 \text{ def in batch} \mid 1 \text{ def in sample}) = \frac{.06}{.06 + .0712} = .457$$

$$P(2 \text{ def in batch} \mid 1 \text{ def in sample}) = \frac{.0712}{.06 + .0712} = .543$$

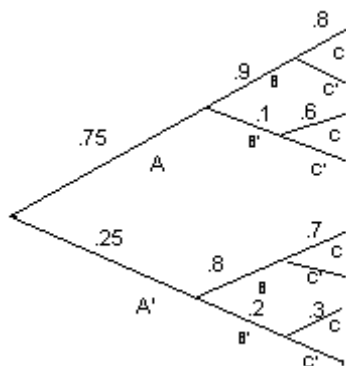
62. Using a tree diagram, B = basic, D = deluxe, W = warranty purchase, W' = no warranty



We want $P(B|W) = \frac{P(B \cap W)}{P(W)} = \frac{.12}{.30 + .12} = \frac{.12}{.42} = .2857$

63.

a.



b. $P(A \cap B \cap C) = .75 \times .9 \times .8 = .5400$

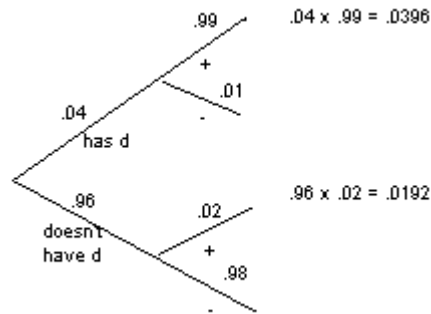
c. $P(B \cap C) = P(A \cap B \cap C) + P(A' \cap B \cap C)$
 $= .5400 + .25 \times .8 \times .7 = .6800$

d. $P(C) = P(A \cap B \cap C) + P(A' \cap B \cap C) + P(A \cap B' \cap C) + P(A' \cap B' \cap C)$
 $= .54 + .045 + .14 + .015 = .74$

e. $P(A|B \cap C) = \frac{P(A \cap B \cap C)}{P(B \cap C)} = \frac{.54}{.68} = .7941$

Chapter 2: Probability

64.

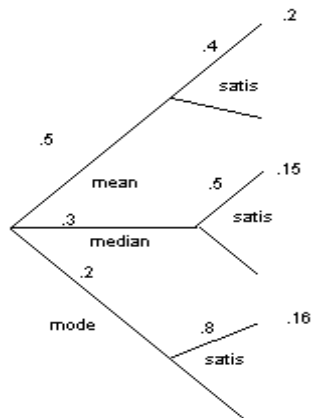


a. $P(+) = .0588$

b. $P(\text{has d} | +) = \frac{.0396}{.0588} = .6735$

c. $P(\text{doesn't have d} | -) = \frac{.9408}{.9412} = .9996$

65.



$P(\text{satis}) = .51$

$P(\text{mean} | \text{satis}) = \frac{.2}{.51} = .3922$

$P(\text{median} | \text{satis}) = .2941$

$P(\text{mode} | \text{satis}) = .3137$

So Mean (and not Mode!) is the most likely author, while Median is least.

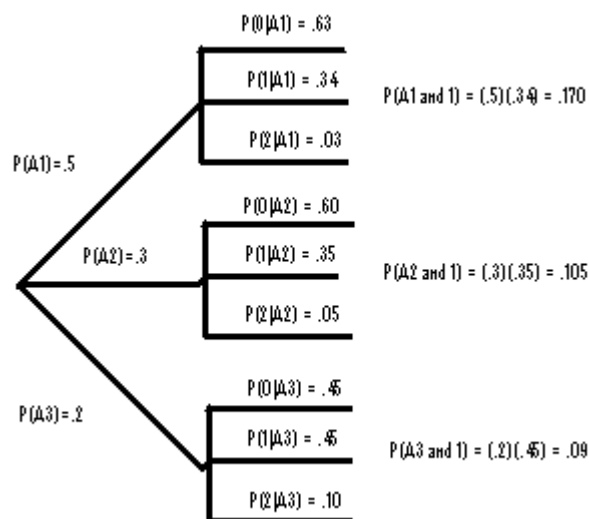
Chapter 2: Probability

- 66.** Define events A_1 , A_2 , and A_3 as flying with airline 1, 2, and 3, respectively. Events 0, 1, and 2 are 0, 1, and 2 flights are late, respectively. Event DC = the event that the flight to DC is late, and event LA = the event that the flight to LA is late. Creating a tree diagram as described in the hint, the probabilities of the second generation branches are calculated as follows: For the A_1 branch, $P(0|A_1) = P[DC' \cap LA'] = P[DC'] \cdot P[LA'] = (.7)(.9) = .63$; $P(1|A_1) = P[(DC' \cap LA) \cup (DC \cap LA')] = (.7)(.1) + (.3)(.9) = .07 + .27 = .34$; $P(2|A_1) = P[DC \cap LA] = P[DC] \cdot P[LA] = (.3)(.1) = .03$. Follow a similar pattern for A_2 and A_3 .

From the law of total probability, we know that

$$\begin{aligned} P(1) &= P(A_1 \cap 1) + P(A_2 \cap 1) + P(A_3 \cap 1) \\ &= (\text{from tree diagram below}) .170 + .105 + .09 = .365. \end{aligned}$$

We wish to find $P(A_1|1)$, $P(A_2|1)$, and $P(A_3|1)$.

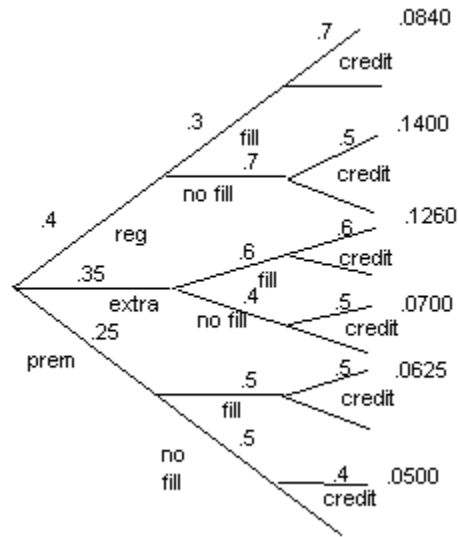


$$P(A_1/1) = \frac{P(A_1 \cap 1)}{P(1)} = \frac{.170}{.365} = .466;$$

$$P(A_2/1) = \frac{P(A_2 \cap 1)}{P(1)} = \frac{.105}{.365} = .288;$$

$$P(A_3/1) = \frac{P(A_3 \cap 1)}{P(1)} = \frac{.090}{.365} = .247;$$

67.



- a. $P(U \cap F \cap Cr) = .1260$
- b. $P(Pr \cap NF \cap Cr) = .05$
- c. $P(Pr \cap Cr) = .0625 + .05 = .1125$
- d. $P(F \cap Cr) = .0840 + .1260 + .0625 = .2725$
- e. $P(Cr) = .5325$
- f. $P(PR | Cr) = \frac{P(Pr \cap Cr)}{P(Cr)} = \frac{.1125}{.5325} = .2113$

Section 2.5

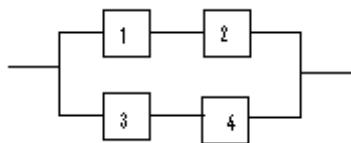
68. Using the definition, two events A and B are independent if $P(A|B) = P(A)$;
 $P(A|B) = .6125$; $P(A) = .50$; $.6125 \neq .50$, so A and B are dependent.
 Using the multiplication rule, the events are independent if
 $P(A \cap B) = P(A) \cdot P(B)$;
 $P(A \cap B) = .25$; $P(A) \cdot P(B) = (.5)(.4) = .2$. $.25 \neq .2$, so A and B are dependent.
- 69.
- a. Since the events are independent, then A' and B' are independent, too. (see paragraph below equation 2.7. $P(B'|A') = .$ $P(B') = 1 - .7 = .3$
- b. $P(A \cup B) = P(A) + P(B) - P(A) \cdot P(B) = .4 + .7 + (.4)(.7) = .82$
- c. $P(AB'|A \cup B) = \frac{P(AB' \cap (A \cup B))}{P(A \cup B)} = \frac{P(AB')}{P(A \cup B)} = \frac{.12}{.82} = .146$
70. $P(A_1 \cap A_2) = .11$, $P(A_1) \cdot P(A_2) = .055$. A_1 and A_2 are not independent.
 $P(A_1 \cap A_3) = .05$, $P(A_1) \cdot P(A_3) = .0616$. A_1 and A_3 are not independent.
 $P(A_2 \cap A_3) = .07$, $P(A_2) \cdot P(A_3) = .07$. A_2 and A_3 are independent.
71. $P(A' \cap B) = P(B) - P(A \cap B) = P(B) - P(A) \cdot P(B) = [1 - P(A)] \cdot P(B) = P(A') \cdot P(B)$.
 Alternatively, $P(A' | B) = \frac{P(A' \cap B)}{P(B)} = \frac{P(B) - P(A \cap B)}{P(B)}$

$$= \frac{P(B) - P(A) \cdot P(B)}{P(B)} = 1 - P(A) = P(A').$$
72. Using subscripts to differentiate between the selected individuals,
 $P(O_1 \cap O_2) = P(O_1) \cdot P(O_2) = (.44)(.44) = .1936$
 $P(\text{two individuals match}) = P(A_1 \cap A_2) + P(B_1 \cap B_2) + P(AB_1 \cap AB_2) + P(O_1 \cap O_2)$
 $= .42^2 + .10^2 + .04^2 + .44^2 = .3816$
73. Let event E be the event that an error was signaled incorrectly. We want $P(\text{at least one signaled incorrectly}) = P(E_1 \cup E_2 \cup \dots \cup E_{10}) = 1 - P(E_1' \cap E_2' \cap \dots \cap E_{10}')$. $P(E') = 1 - .05 = .95$. For 10 independent points, $P(E_1' \cap E_2' \cap \dots \cap E_{10}') = P(E_1')P(E_2') \dots P(E_{10}')$ so $= P(E_1 \cup E_2 \cup \dots \cup E_{10}) = 1 - [.95]^{10} = .401$. Similarly, for 25 points, the desired probability is $= 1 - [P(E')^{25}] = 1 - (.95)^{25} = .723$

Chapter 2: Probability

74. $P(\text{no error on any particular question}) = .9$, so $P(\text{no error on any of the 10 questions}) = (.9)^{10} = .3487$. Then $P(\text{at least one error}) = 1 - (.9)^{10} = .6513$. For p replacing .1, the two probabilities are $(1-p)^n$ and $1 - (1-p)^n$.
75. Let q denote the probability that a rivet is defective.
- a. $P(\text{seam need rework}) = .20 = 1 - P(\text{seam doesn't need rework})$
 $= 1 - P(\text{no rivets are defective})$
 $= 1 - P(1^{\text{st}} \text{ isn't def} \cap \dots \cap 25^{\text{th}} \text{ isn't def})$
 $= 1 - (1 - q)^{25}$, so $.80 = (1 - q)^{25}$, $1 - q = (.80)^{1/25}$, and thus $q = 1 - .99111 = .00889$.
- b. The desired condition is $.10 = 1 - (1 - q)^{25}$, i.e. $(1 - q)^{25} = .90$, from which $q = 1 - .99579 = .00421$.
76. $P(\text{at least one opens}) = 1 - P(\text{none open}) = 1 - (.05)^5 = .99999969$
 $P(\text{at least one fails to open}) = 1 - P(\text{all open}) = 1 - (.95)^5 = .2262$
77. Let A_1 = older pump fails, A_2 = newer pump fails, and $x = P(A_1 \cap A_2)$. Then $P(A_1) = .10 + x$, $P(A_2) = .05 + x$, and $x = P(A_1 \cap A_2) = P(A_1) \cdot P(A_2) = (.10 + x)(.05 + x)$. The resulting quadratic equation, $x^2 - .85x + .005 = 0$, has roots $x = .0059$ and $x = .8441$. Hopefully the smaller root is the actual probability of system failure.
78. $P(\text{system works}) = P(1 - 2 \text{ works} \cup 3 - 4 \text{ works})$
 $= P(1 - 2 \text{ works}) + P(3 - 4 \text{ works}) - P(1 - 2 \text{ works} \cap 3 - 4 \text{ works})$
 $= P(1 \text{ works} \cup 2 \text{ works}) + P(3 \text{ works} \cap 4 \text{ works}) - P(1 - 2) \cdot P(3 - 4)$
 $= (.9 + .9 \cdot .81) + (.9)(.9) - (.9 + .9 \cdot .81)(.9)(.9)$
 $= .99 + .81 - .8019 = .9981$

79.



Using the hints, let $P(A_i) = p$, and $x = p^2$, then $P(\text{system lifetime exceeds } t_0) = p^2 + p^2 - p^4 = 2p^2 - p^4 = 2x - x^2$. Now, set this equal to .99, or $2x - x^2 = .99 \Rightarrow x^2 - 2x + .99 = 0$. Use the

quadratic formula to solve for x :
$$= \frac{2 \pm \sqrt{4 - (4)(.99)}}{2} = \frac{2 \pm .2}{2} = 1 \pm .1 = .99 \text{ or } 1.01$$

Since the value we want is a probability, and has to be ≤ 1 , we use the value of .99.

80. Event A: $\{ (3,1)(3,2)(3,3)(3,4)(3,5)(3,6) \}$, $P(A) = \frac{1}{6}$;

Event B: $\{ (1,4)(2,4)(3,4)(4,4)(5,4)(6,4) \}$, $P(B) = \frac{1}{6}$;

Event C: $\{ (1,6)(2,5)(3,4)(4,3)(5,2)(6,1) \}$, $P(C) = \frac{1}{6}$;

Event $A \cap B$: $\{ (3,4) \}$; $P(A \cap B) = \frac{1}{36}$;

Event $A \cap C$: $\{ (3,4) \}$; $P(A \cap C) = \frac{1}{36}$;

Event $B \cap C$: $\{ (3,4) \}$; $P(B \cap C) = \frac{1}{36}$;

Event $A \cap B \cap C$: $\{ (3,4) \}$; $P(A \cap B \cap C) = \frac{1}{36}$;

$$P(A) \cdot P(B) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36} = P(A \cap B)$$

$$P(A) \cdot P(C) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36} = P(A \cap C)$$

$$P(B) \cdot P(C) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36} = P(B \cap C)$$

The events are pairwise independent.

$$P(A) \cdot P(B) \cdot P(C) = \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{216} \neq \frac{1}{36} = P(A \cap B \cap C)$$

The events are not mutually independent

Chapter 2: Probability

- 81.** $P(\text{both detect the defect}) = 1 - P(\text{at least one doesn't}) = 1 - .2 = .8$
- a.** $P(1^{\text{st}} \text{ detects} \cap 2^{\text{nd}} \text{ doesn't}) = P(1^{\text{st}} \text{ detects}) - P(1^{\text{st}} \text{ does} \cap 2^{\text{nd}} \text{ does})$
 $= .9 - .8 = .1$
 Similarly, $P(1^{\text{st}} \text{ doesn't} \cap 2^{\text{nd}} \text{ does}) = .1$, so $P(\text{exactly one does}) = .1 + .1 = .2$
- b.** $P(\text{neither detects a defect}) = 1 - [P(\text{both do}) + P(\text{exactly 1 does})]$
 $= 1 - [.8 + .2] = 0$
 so $P(\text{all 3 escape}) = (0)(0)(0) = 0$.
- 82.** $P(\text{pass}) = .70$
- a.** $(.70)(.70)(.70) = .343$
- b.** $1 - P(\text{all pass}) = 1 - .343 = .657$
- c.** $P(\text{exactly one passes}) = (.70)(.30)(.30) + (.30)(.70)(.30) + (.30)(.30)(.70) = .189$
- d.** $P(\# \text{ pass} \leq 1) = P(0 \text{ pass}) + P(\text{exactly one passes}) = (.3)^3 + .189 = .216$
- e.** $P(3 \text{ pass} \mid 1 \text{ or more pass}) =$
 $= \frac{P(3 \text{ pass} \cap \geq 1 \text{ pass})}{P(\geq 1 \text{ pass})} = \frac{P(3 \text{ pass})}{P(\geq 1 \text{ pass})} = \frac{.343}{.973} = .353$
- 83.**
- a.** Let D_1 = detection on 1st fixation, D_2 = detection on 2nd fixation.
 $P(\text{detection in at most 2 fixations}) = P(D_1) + P(D_1' \cap D_2)$
 $= P(D_1) + P(D_2 \mid D_1')P(D_1')$
 $= p + p(1 - p) = p(2 - p)$.
- b.** Define $D_1, D_2, \dots, D_n$ as in **a**. Then $P(\text{at most } n \text{ fixations})$
 $= P(D_1) + P(D_1' \cap D_2) + P(D_1' \cap D_2' \cap D_3) + \dots + P(D_1' \cap D_2' \cap \dots \cap D_{n-1}' \cap D_n)$
 $= p + p(1 - p) + p(1 - p)^2 + \dots + p(1 - p)^{n-1}$
 $= p [1 + (1 - p) + (1 - p)^2 + \dots + (1 - p)^{n-1}] = p \bullet \frac{1 - (1 - p)^n}{1 - (1 - p)} = 1 - (1 - p)^n$
 Alternatively, $P(\text{at most } n \text{ fixations}) = 1 - P(\text{at least } n+1 \text{ are req'd})$
 $= 1 - P(\text{no detection in 1^{st } n \text{ fixations})}$
 $= 1 - P(D_1' \cap D_2' \cap \dots \cap D_n')$
 $= 1 - (1 - p)^n$
- c.** $P(\text{no detection in 3 fixations}) = (1 - p)^3$

Chapter 2: Probability

$$\begin{aligned}
 \text{d. } P(\text{passes inspection}) &= P(\{\text{not flawed}\} \cup \{\text{flawed and passes}\}) \\
 &= P(\text{not flawed}) + P(\text{flawed and passes}) \\
 &= .9 + P(\text{passes} \mid \text{flawed}) \cdot P(\text{flawed}) = .9 + (1 - p)^3(.1)
 \end{aligned}$$

$$\text{e. } P(\text{flawed} \mid \text{passed}) = \frac{P(\text{flawed} \cap \text{passed})}{P(\text{passed})} = \frac{.1(1 - p)^3}{.9 + .1(1 - p)^3}$$

$$\text{For } p = .5, P(\text{flawed} \mid \text{passed}) = \frac{.1(.5)^3}{.9 + .1(.5)^3} = .0137$$

84.

$$\begin{aligned}
 \text{a. } P(A) &= \frac{2000}{10,000} = .02, P(B) = P(A \cap B) + P(A' \cap B) \\
 &= P(B \mid A) P(A) + P(B \mid A') P(A') = \frac{1999}{9999} \cdot (.2) + \frac{2000}{9999} \cdot (.8) = .2 \\
 P(A \cap B) &= .039984; \text{ since } P(A \cap B) \neq P(A)P(B), \text{ the events are not independent.}
 \end{aligned}$$

$$\text{b. } P(A \cap B) = .04. \text{ Very little difference. Yes.}$$

$$\begin{aligned}
 \text{c. } P(A) &= P(B) = .2, P(A)P(B) = .04, \text{ but } P(A \cap B) = P(B \mid A)P(A) = \frac{1}{9} \cdot \frac{2}{10} = .0222, \text{ so the} \\
 &\text{two numbers are quite different.} \\
 \text{In a, the sample size is small relative to the "population" size, while here it is not.}
 \end{aligned}$$

85.

$$\begin{aligned}
 P(\text{system works}) &= P(1 - 2 \text{ works} \cap 3 - 4 - 5 - 6 \text{ works} \cap 7 \text{ works}) \\
 &= P(1 - 2 \text{ works}) \cdot P(3 - 4 - 5 - 6 \text{ works}) \cdot P(7 \text{ works}) \\
 &= (.99) (.9639) (.9) = .8588
 \end{aligned}$$

$$\begin{aligned}
 &\text{With the subsystem in figure 2.14 connected in parallel to this subsystem,} \\
 P(\text{system works}) &= .8588 + .927 - (.8588)(.927) = .9897
 \end{aligned}$$

Chapter 2: Probability

86.

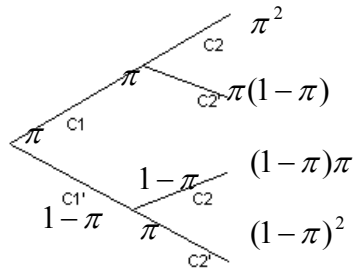
- a. For route #1, $P(\text{late}) = P(\text{stopped at 2 or 3 or 4 crossings})$
 $= 1 - P(\text{stopped at 0 or 1}) = 1 - [.9^4 + 4(.9)^3(.1)]$
 $= .0523$

For route #2, $P(\text{late}) = P(\text{stopped at 1 or 2 crossings})$
 $= 1 - P(\text{stopped at none}) = 1 - .81 = .19$
 thus route #1 should be taken.

b. $P(4 \text{ crossing route} \mid \text{late}) = \frac{P(4 \text{ crossing route} \cap \text{late})}{P(\text{late})}$

$$= \frac{(.5)(.0523)}{(.5)(.0523) + (.5)(.19)} = .216$$

87.



$$P(\text{at most 1 is lost}) = 1 - P(\text{both lost})$$

$$= 1 - \pi^2$$

$$P(\text{exactly 1 lost}) = 2\pi(1 - \pi)$$

$$P(\text{exactly 1} \mid \text{at most 1}) = \frac{P(\text{exactly 1})}{P(\text{at most 1})} = \frac{2\pi(1 - \pi)}{1 - \pi^2}$$

Supplementary Exercises

88.

a. $\binom{20}{3} = 1140$

b. $\binom{19}{3} = 969$

c. # having at least 1 of the 10 best = $1140 - \# \text{ of crews having none of 10 best} = 1140 -$

$$\binom{10}{3} = 1140 - 120 = 1020$$

d. $P(\text{best will not work}) = \frac{969}{1140} = .85$

89.

a. $P(\text{line 1}) = \frac{500}{1500} = .333;$

$$P(\text{Crack}) = \frac{.50(500) + .44(400) + .40(600)}{1500} = \frac{666}{1500} = .444$$

b. $P(\text{Blemish} \mid \text{line 1}) = .15$

c. $P(\text{Surface Defect}) = \frac{.10(500) + .08(400) + .15(600)}{1500} = \frac{172}{1500}$

$$P(\text{line 1 and Surface Defect}) = \frac{.10(500)}{1500} = \frac{50}{1500}$$

$$\text{So } P(\text{line 1} \mid \text{Surface Defect}) = \frac{50/1500}{172/1500} = .291$$

90.

- a. The only way he will have one type of forms left is if they are all course substitution forms. He must choose all 6 of the withdrawal forms to pass to a subordinate. The

desired probability is $\frac{\binom{6}{6}}{\binom{10}{6}} = .00476$

- b. He can start with the wd forms: W-C-W-C or with the cs forms: C-W-C-W:

of ways: $6 \times 4 \times 5 \times 3 + 4 \times 6 \times 3 \times 5 = 2(360) = 720;$

The total # ways to arrange the four forms: $10 \times 9 \times 8 \times 7 = 5040.$

The desired probability is $720/5040 = .1429$

Chapter 2: Probability

91. $P(A \cup B) = P(A) + P(B) - P(A)P(B)$
 $.626 = P(A) + P(B) - .144$

So $P(A) + P(B) = .770$ and $P(A)P(B) = .144$.

Let $x = P(A)$ and $y = P(B)$, then using the first equation, $y = .77 - x$, and substituting this into the second equation, we get $x(.77 - x) = .144$ or $x^2 - .77x + .144 = 0$. Use the quadratic formula to solve:

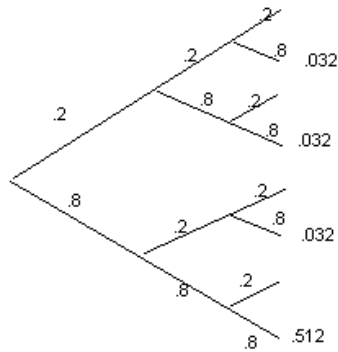
$$\frac{.77 \pm \sqrt{.77^2 - (4)(.144)}}{2} = \frac{.77 \pm .13}{2} = .32 \text{ or } .45$$

So $P(A) = .45$ and $P(B) = .32$

92.

a. $(.8)(.8)(.8) = .512$

b.



$$.512 + .032 + .032 + .032 = .608$$

c. $P(1 \text{ sent} | 1 \text{ received}) = \frac{P(1 \text{ sent} \cap 1 \text{ received})}{P(1 \text{ received})} = \frac{.4256}{.5432} = .7835$

Chapter 2: Probability

93.

a. There are $5 \times 4 \times 3 \times 2 \times 1 = 120$ possible orderings, so $P(BCDEF) = \frac{1}{120} = .0083$

b. # orderings in which F is 3rd = $4 \times 3 \times 1 * \times 2 \times 1 = 24$, (* because F must be here), so
 $P(F \text{ 3}^{\text{rd}}) = \frac{24}{120} = .2$

c. $P(F \text{ last}) = \frac{4 \times 3 \times 2 \times 1 \times 1}{120} = .2$

94. $P(F \text{ hasn't heard after 10 times}) = P(\text{not on \#1} \cap \text{not on \#2} \cap \dots \cap \text{not on \#10})$

$$= \left(\frac{4}{5}\right)^{10} = .1074$$

95. When three experiments are performed, there are 3 different ways in which detection can occur on exactly 2 of the experiments: (i) #1 and #2 and not #3 (ii) #1 and not #2 and #3; (iii) not #1 and #2 and #3. If the impurity is present, the probability of exactly 2 detections in three (independent) experiments is $(.8)(.8)(.2) + (.8)(.2)(.8) + (.2)(.8)(.8) = .384$. If the impurity is absent, the analogous probability is $3(.1)(.1)(.9) = .027$. Thus

$P(\text{present} \mid \text{detected in exactly 2 out of 3}) =$

$$\frac{P(\text{detected in exactly 2} \cap \text{present})}{P(\text{detected in exactly 2})}$$

$$= \frac{(.384)(.4)}{(.384)(.4) + (.027)(.6)} = .905$$

96. $P(\text{exactly 1 selects category \#1} \mid \text{all 3 are different})$

$$= \frac{P(\text{exactly 1 selects \#1} \cap \text{all are different})}{P(\text{all are different})}$$

$$\text{Denominator} = \frac{6 \times 5 \times 4}{6 \times 6 \times 6} = \frac{5}{9} = .5556$$

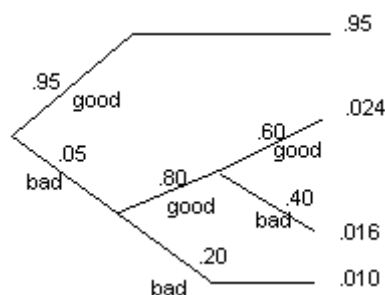
Numerator = 3 $P(\text{contestant \#1 selects category \#1 and the other two select two different categories})$

$$= 3 \times \frac{1 \times 5 \times 4}{6 \times 6 \times 6} = \frac{5 \times 4 \times 3}{6 \times 6 \times 6}$$

$$\text{The desired probability is then } \frac{5 \times 4 \times 3}{6 \times 5 \times 4} = \frac{1}{2} = .5$$

Chapter 2: Probability

97.



a. $P(\text{pass inspection}) = P(\text{pass initially} \cup \text{passes after recrimping}) = P(\text{pass initially}) + P(\text{fails initially} \cap \text{goes to recrimping} \cap \text{is corrected after recrimping})$
 $= .95 + (.05)(.80)(.60)$ (following path “bad-good-good” on tree diagram)
 $= .974$

b. $P(\text{needed no recrimping} \mid \text{passed inspection}) = \frac{P(\text{passed initially})}{P(\text{passed inspection})}$
 $= \frac{.95}{.974} = .9754$

98.

a. $P(\text{both} +) = P(\text{carrier} \cap \text{both} +) + P(\text{not a carrier} \cap \text{both} +)$
 $= P(\text{both} + \mid \text{carrier}) \times P(\text{carrier})$
 $+ P(\text{both} + \mid \text{not a carrier}) \times P(\text{not a carrier})$
 $= (.90)^2(.01) + (.05)^2(.99) = .01058$
 $P(\text{both} -) = (.10)^2(.01) + (.95)^2(.99) = .89358$
 $P(\text{tests agree}) = .01058 + .89358 = .90416$

b. $P(\text{carrier} \mid \text{both} + \text{ve}) = \frac{P(\text{carrier} \cap \text{both. positive})}{P(\text{both. positive})} = \frac{(.90)^2(.01)}{.01058} = .7656$

99. Let $A = 1^{\text{st}}$ functions, $B = 2^{\text{nd}}$ functions, so $P(B) = .9$, $P(A \cup B) = .96$, $P(A \cap B) = .75$. Thus, $P(A \cup B) = P(A) + P(B) - P(A \cap B) = P(A) + .9 - .75 = .96$, implying $P(A) = .81$.

This gives $P(B \mid A) = \frac{P(B \cap A)}{P(A)} = \frac{.75}{.81} = .926$

100. $P(E_1 \cap \text{late}) = P(\text{late} \mid E_1)P(E_1) = (.02)(.40) = .008$

Chapter 2: Probability

101.

- a. The law of total probability gives

$$\begin{aligned} P(\text{late}) &= \sum_{i=1}^3 P(\text{late} | E_i) \cdot P(E_i) \\ &= (.02)(.40) + (.01)(.50) + (.05)(.10) = .018 \end{aligned}$$

- b. $P(E_1' | \text{on time}) = 1 - P(E_1 | \text{on time})$

$$= 1 - \frac{P(E_1 \cap \text{on.time})}{P(\text{on.time})} = 1 - \frac{(.98)(.4)}{.982} = .601$$

102. Let B denote the event that a component needs rework. Then

$$P(B) = \sum_{i=1}^3 P(B | A_i) \cdot P(A_i) = (.05)(.50) + (.08)(.30) + (.10)(.20) = .069$$

$$\text{Thus } P(A_1 | B) = \frac{(.05)(.50)}{.069} = .362$$

$$P(A_2 | B) = \frac{(.08)(.30)}{.069} = .348$$

$$P(A_3 | B) = \frac{(.10)(.20)}{.069} = .290$$

103.

a. $P(\text{all different}) = \frac{(365)(364)\dots(356)}{(365)^{10}} = .883$

$$P(\text{at least two the same}) = 1 - .883 = .117$$

- b. $P(\text{at least two the same}) = .476$ for $k=22$, and $= .507$ for $k=23$

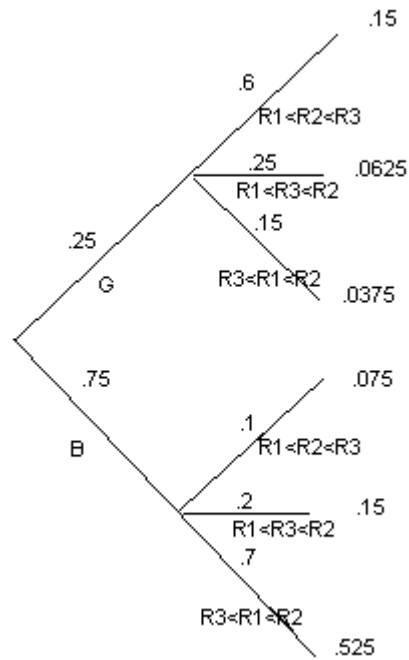
- c. $P(\text{at least two have the same SS number}) = 1 - P(\text{all different})$

$$= 1 - \frac{(1000)(999)\dots(991)}{(1000)^{10}}$$

$$= 1 - .956 = .044$$

$$\begin{aligned} \text{Thus } P(\text{at least one "coincidence"}) &= P(\text{BD coincidence} \cup \text{SS coincidence}) \\ &= .117 + .044 - (.117)(.044) = .156 \end{aligned}$$

104.



a. $P(G | R_1 < R_2 < R_3) = \frac{.15}{.15 + .075} = .67$, $P(B | R_1 < R_2 < R_3) = .33$, classify as granite.

b. $P(G | R_1 < R_3 < R_2) = \frac{.0625}{.2125} = .2941 < .05$, so classify as basalt.

$$P(G | R_3 < R_1 < R_2) = \frac{.0375}{.5625} = .0667, \text{ so classify as basalt.}$$

c. $P(\text{erroneous classif}) = P(B \text{ classif as } G) + P(G \text{ classif as } B)$
 $= P(\text{classif as } G | B)P(B) + P(\text{classif as } B | G)P(G)$
 $= P(R_1 < R_2 < R_3 | B)(.75) + P(R_1 < R_3 < R_2 \text{ or } R_3 < R_1 < R_2 | G)(.25)$
 $= (.10)(.75) + (.25 + .15)(.25) = .175$

Chapter 2: Probability

- d. For what values of p will $P(G \mid R_1 < R_2 < R_3) > .5$, $P(G \mid R_1 < R_3 < R_2) > .5$, $P(G \mid R_3 < R_1 < R_2) > .5$?

$$P(G \mid R_1 < R_2 < R_3) = \frac{.6p}{.6p + .1(1-p)} = \frac{.6p}{.1 + .5p} > .5 \text{ iff } p > \frac{1}{7}$$

$$P(G \mid R_1 < R_3 < R_2) = \frac{.25p}{.25p + .2(1-p)} > .5 \text{ iff } p > \frac{4}{9}$$

$$P(G \mid R_3 < R_1 < R_2) = \frac{.15p}{.15p + .7(1-p)} > .5 \text{ iff } p > \frac{14}{17} \text{ (most restrictive)}$$

If $p > \frac{14}{17}$ always classify as granite.

105. $P(\text{detection by the end of the } n\text{th glimpse}) = 1 - P(\text{not detected in } 1^{\text{st}} n)$

$$= 1 - P(G_1' \cap G_2' \cap \dots \cap G_n') = 1 - P(G_1')P(G_2') \dots P(G_n')$$

$$= 1 - (1 - p_1)(1 - p_2) \dots (1 - p_n) = 1 - \prod_{i=1}^n (1 - p_i)$$

106.

a. $P(\text{walks on } 4^{\text{th}} \text{ pitch}) = P(1^{\text{st}} 4 \text{ pitches are balls}) = (.5)^4 = .0625$

b. $P(\text{walks on } 6^{\text{th}}) = P(2 \text{ of the } 1^{\text{st}} 5 \text{ are strikes, \#6 is a ball})$
 $= P(2 \text{ of the } 1^{\text{st}} 5 \text{ are strikes})P(\#6 \text{ is a ball})$
 $= [10(.5)^5](.5) = .15625$

c. $P(\text{Batter walks}) = P(\text{walks on } 4^{\text{th}}) + P(\text{walks on } 5^{\text{th}}) + P(\text{walks on } 6^{\text{th}})$
 $= .0625 + .15625 + .15625 = .375$

d. $P(\text{first batter scores while no one is out}) = P(\text{first 4 batters walk})$
 $= (.375)^4 = .0198$

107.

a. $P(\text{all in correct room}) = \frac{1}{4 \times 3 \times 2 \times 1} = \frac{1}{24} = .0417$

b. The 9 outcomes which yield incorrect assignments are: 2143, 2341, 2413, 3142, 3412, 3421, 4123, 4321, and 4312, so $P(\text{all incorrect}) = \frac{9}{24} = .375$

Chapter 2: Probability

108.

- a. $P(\text{all full}) = P(A \cap B \cap C) = (.6)(.5)(.4) = .12$
 $P(\text{at least one isn't full}) = 1 - P(\text{all full}) = 1 - .12 = .88$
- b. $P(\text{only NY is full}) = P(A \cap B' \cap C') = P(A)P(B')P(C') = .18$
 Similarly, $P(\text{only Atlanta is full}) = .12$ and $P(\text{only LA is full}) = .08$
 So $P(\text{exactly one full}) = .18 + .12 + .08 = .38$

109. Note: $s = 0$ means that the very first candidate interviewed is hired. Each entry below is the candidate hired for the given policy and outcome.

Outcome	s=0	s=1	s=2	s=3	Outcome	s=0	s=1	s=2	s=3
1234	1	4	4	4	3124	3	1	4	4
1243	1	3	3	3	3142	3	1	4	2
1324	1	4	4	4	3214	3	2	1	4
1342	1	2	2	2	3241	3	2	1	1
1423	1	3	3	3	3412	3	1	1	2
1432	1	2	2	2	3421	3	2	2	1
2134	2	1	4	4	4123	4	1	3	3
2143	2	1	3	3	4132	4	1	2	2
2314	2	1	1	4	4213	4	2	1	3
2341	2	1	1	1	4231	4	2	1	1
2413	2	1	1	3	4312	4	3	1	2
2431	2	1	1	1	4321	4	3	2	1

s	0	1	2	3
P(hire#1)	$\frac{6}{24}$	$\frac{11}{24}$	$\frac{10}{24}$	$\frac{6}{24}$

So $s = 1$ is best.

110. $P(\text{at least one occurs}) = 1 - P(\text{none occur})$
 $= 1 - (1 - p_1)(1 - p_2)(1 - p_3)(1 - p_4)$
 $= p_1p_2(1 - p_3)(1 - p_4) + \dots + (1 - p_1)(1 - p_2)p_3p_4$
 $+ (1 - p_1)p_2p_3p_4 + \dots + p_1p_2p_3(1 - p_4) + p_1p_2p_3p_4$

111. $P(A_1) = P(\text{draw slip 1 or 4}) = \frac{1}{2}$; $P(A_2) = P(\text{draw slip 2 or 4}) = \frac{1}{2}$;
 $P(A_3) = P(\text{draw slip 3 or 4}) = \frac{1}{2}$; $P(A_1 \cap A_2) = P(\text{draw slip 4}) = \frac{1}{4}$;
 $P(A_2 \cap A_3) = P(\text{draw slip 4}) = \frac{1}{4}$; $P(A_1 \cap A_3) = P(\text{draw slip 4}) = \frac{1}{4}$
 Hence $P(A_1 \cap A_2) = P(A_1)P(A_2) = \frac{1}{4}$, $P(A_2 \cap A_3) = P(A_2)P(A_3) = \frac{1}{4}$,
 $P(A_1 \cap A_3) = P(A_1)P(A_3) = \frac{1}{4}$, thus there exists pairwise independence
 $P(A_1 \cap A_2 \cap A_3) = P(\text{draw slip 4}) = \frac{1}{4} \neq \frac{1}{8} = P(A_1)P(A_2)P(A_3)$, so the events are not mutually independent.

CHAPTER 3

Section 3.1

1.

S:	FFF	SFF	FSF	FFS	FSS	SFS	SSF	SSS
X:	0	1	1	1	2	2	2	3

2.

$X = 1$ if a randomly selected book is non-fiction and $X = 0$ otherwise
 $X = 1$ if a randomly selected executive is a female and $X = 0$ otherwise
 $X = 1$ if a randomly selected driver has automobile insurance and $X = 0$ otherwise

3.

M = the difference between the large and the smaller outcome with possible values 0, 1, 2, 3, 4, or 5; $W = 1$ if the sum of the two resulting numbers is even and $W = 0$ otherwise, a Bernoulli random variable.

4.

In my perusal of a zip code directory, I found no 00000, nor did I find any zip codes with four zeros, a fact which was not obvious. Thus possible X values are 2, 3, 4, 5 (and not 0 or 1). $X = 5$ for the outcome 15213, $X = 4$ for the outcome 44074, and $X = 3$ for 94322.

5.

No. In the experiment in which a coin is tossed repeatedly until a H results, let $Y = 1$ if the experiment terminates with at most 5 tosses and $Y = 0$ otherwise. The sample space is infinite, yet Y has only two possible values.

6.

Possible X values are 1, 2, 3, 4, ... (all positive integers)

Outcome:	RL	AL	RAARL	RRRRL	AARRL
X:	2	2	5	5	5

Chapter 3: Discrete Random Variables and Probability Distributions

7.

- a. Possible values are $0, 1, 2, \dots, 12$; discrete
- b. With $N = \#$ on the list, values are $0, 1, 2, \dots, N$; discrete
- c. Possible values are $1, 2, 3, 4, \dots$; discrete
- d. $\{x: 0 < x < \infty\}$ if we assume that a rattlesnake can be arbitrarily short or long; not discrete
- e. With $c =$ amount earned per book sold, possible values are $0, c, 2c, 3c, \dots, 10,000c$; discrete
- f. $\{y: 0 < y < 14\}$ since 0 is the smallest possible pH and 14 is the largest possible pH; not discrete
- g. With m and M denoting the minimum and maximum possible tension, respectively, possible values are $\{x: m < x < M\}$; not discrete
- h. Possible values are $3, 6, 9, 12, 15, \dots$ -- i.e. $3(1), 3(2), 3(3), 3(4), \dots$ giving a first element, etc.; discrete

8.

$Y = 3$: SSS; $Y = 4$: FSSS; $Y = 5$: FFSSS, SFSSS;
 $Y = 6$: SSFSSS, SFFSSS, FSFSSS, FFFSSS;
 $Y = 7$: SSFFS, SFSFSSS, SFFFSSS, FSSFSSS, FSFFSSS, FFSFSSS, FFFFSSS

9.

- a. Returns to 0 can occur only after an even number of tosses; possible S values are $2, 4, 6, 8, \dots$ (i.e. $2(1), 2(2), 2(3), 2(4), \dots$) an infinite sequence, so x is discrete.
- b. Now a return to 0 is possible after any number of tosses greater than 1, so possible values are $2, 3, 4, 5, \dots$ ($1+1, 1+2, 1+3, 1+4, \dots$, an infinite sequence) and X is discrete

10.

- a. $T =$ total number of pumps in use at both stations. Possible values: $0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10$
- b. X : $-4, -3, -2, -1, 0, 1, 2, 3, 4, 5, 6$
- c. U : $0, 1, 2, 3, 4, 5, 6$
- d. Z : $0, 1, 2$

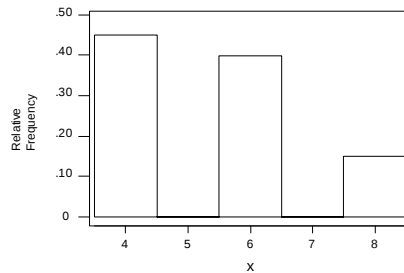
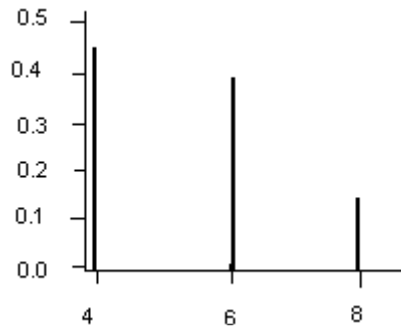
Section 3.2

11.

a.

x	4	6	8
P(x)	.45	.40	.15

b.



c. $P(x \geq 6) = .40 + .15 = .55$ $P(x > 6) = .15$

12.

- a. In order for the flight to accommodate all the ticketed passengers who show up, no more than 50 can show up. We need $y \leq 50$.
 $P(y \leq 50) = .05 + .10 + .12 + .14 + .25 + .17 = .83$
- b. Using the information in a. above, $P(y > 50) = 1 - P(y \leq 50) = 1 - .83 = .17$
- c. For you to get on the flight, at most 49 of the ticketed passengers must show up. $P(y \leq 49) = .05 + .10 + .12 + .14 + .25 = .66$. For the 3rd person on the standby list, at most 47 of the ticketed passengers must show up. $P(y \leq 44) = .05 + .10 + .12 = .27$

Chapter 3: Discrete Random Variables and Probability Distributions

13.

- a. $P(X \leq 3) = p(0) + p(1) + p(2) + p(3) = .10 + .15 + .20 + .25 = .70$
- b. $P(X < 3) = P(X \leq 2) = p(0) + p(1) + p(2) = .45$
- c. $P(3 \leq X) = p(3) + p(4) + p(5) + p(6) = .55$
- d. $P(2 \leq X \leq 5) = p(2) + p(3) + p(4) + p(5) = .71$
- e. The number of lines not in use is $6 - X$, so $6 - X = 2$ is equivalent to $X = 4$, $6 - X = 3$ to $X = 3$, and $6 - X = 4$ to $X = 2$. Thus we desire $P(2 \leq X \leq 4) = p(2) + p(3) + p(4) = .65$
- f. $6 - X \geq 4$ if $6 - 4 \geq X$, i.e. $2 \geq X$, or $X \leq 2$, and $P(X \leq 2) = .10 + .15 + .20 = .45$

14.

- a. $\sum_{y=1}^5 p(y) = K[1 + 2 + 3 + 4 + 5] = 15K = 1 \Rightarrow K = \frac{1}{15}$
- b. $P(Y \leq 3) = p(1) + p(2) + p(3) = \frac{6}{15} = .4$
- c. $P(2 \leq Y \leq 4) = p(2) + p(3) + p(4) = \frac{9}{15} = .6$
- d. $\sum_{y=1}^5 \left(\frac{y^2}{50} \right) = \frac{1}{50}[1 + 4 + 9 + 16 + 25] = \frac{55}{50} \neq 1$; No

15.

- a. (1,2) (1,3) (1,4) (1,5) (2,3) (2,4) (2,5) (3,4) (3,5) (4,5)
- b. $P(X = 0) = p(0) = P[\{(3,4) (3,5) (4,5)\}] = \frac{3}{10} = .3$
 $P(X = 2) = p(2) = P[\{(1,2)\}] = \frac{1}{10} = .1$
 $P(X = 1) = p(1) = 1 - [p(0) + p(2)] = .60$, and $p(x) = 0$ if $x \neq 0, 1, 2$
- c. $F(0) = P(X \leq 0) = P(X = 0) = .30$
 $F(1) = P(X \leq 1) = P(X = 0 \text{ or } 1) = .90$
 $F(2) = P(X \leq 2) = 1$

The c.d.f. is

$$F(x) = \begin{cases} 0 & x < 0 \\ .30 & 0 \leq x < 1 \\ .90 & 1 \leq x < 2 \\ 1 & 2 \leq x \end{cases}$$

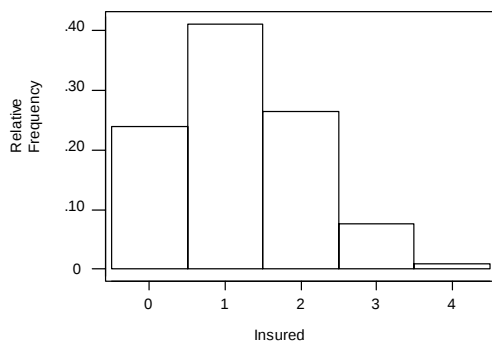
Chapter 3: Discrete Random Variables and Probability Distributions

16.

a.

x	Outcomes		p(x)
0	FFFF	$(.7)^4$	=.2401
1	FFFS,FFSF,FSFF,SFFF	$4[(.7)^3(.3)]$	=.4116
2	FFSS,FSFS,SFFS,FSSF,SFSF,SSFF	$6[(.7)^2(.3)^2]$	=.2646
3	FSSS, SFSS,SSFS,SSSF	$4[(.7)(.3)^3]$	=.0756
4	SSSS	$(.3)^4$	=.0081

b.



c. $p(x)$ is largest for $X = 1$

d. $P(X \geq 2) = p(2) + p(3) + p(4) = .2646 + .0756 + .0081 = .3483$
This could also be done using the complement.

17.

a. $P(2) = P(Y = 2) = P(1^{\text{st}} 2 \text{ batteries are acceptable})$
 $= P(AA) = (.9)(.9) = .81$

b. $p(3) = P(Y = 3) = P(UAA \text{ or } AUA) = (.1)(.9)^2 + (.1)(.9)^2 = 2[(.1)(.9)^2] = .162$

c. The fifth battery must be an A, and one of the first four must also be an A. Thus, $p(5) = P(AUUUA \text{ or } UAUA \text{ or } UUAUA \text{ or } UUUA) = 4[(.1)^3(.9)^2] = .00324$

d. $P(Y = y) = p(y) = P(\text{the } y^{\text{th}} \text{ is an A and so is exactly one of the first } y - 1)$
 $= (y - 1)(.1)^{y-2}(.9)^2, y = 2, 3, 4, 5, \dots$

Chapter 3: Discrete Random Variables and Probability Distributions

18.

a. $p(1) = P(M = 1) = P[(1,1)] = \frac{1}{36}$

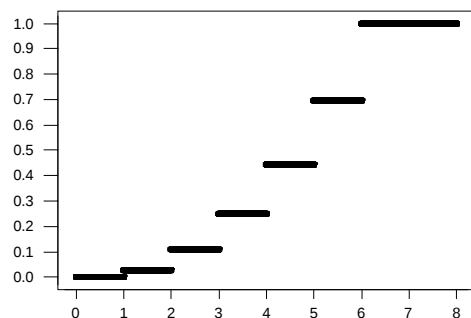
$$p(2) = P(M = 2) = P[(1,2) \text{ or } (2,1) \text{ or } (2,2)] = \frac{3}{36}$$

$$p(3) = P(M = 3) = P[(1,3) \text{ or } (2,3) \text{ or } (3,1) \text{ or } (3,2) \text{ or } (3,3)] = \frac{5}{36}$$

$$\text{Similarly, } p(4) = \frac{7}{36}, p(5) = \frac{9}{36}, \text{ and } p(6) = \frac{11}{36}$$

b. $F(m) = \begin{cases} 0 & \text{for } m < 1, \\ \frac{1}{36} & \text{for } 1 \leq m < 2, \end{cases}$

$$F(m) = \begin{cases} 0 & m < 1 \\ \frac{1}{36} & 1 \leq m < 2 \\ \frac{4}{36} & 2 \leq m < 3 \\ \frac{9}{36} & 3 \leq m < 4 \\ \frac{16}{36} & 4 \leq m < 5 \\ \frac{25}{36} & 5 \leq m < 6 \\ 1 & m \geq 6 \end{cases}$$



19. Let A denote the type O+ individual (type O positive blood) and B, C, D, the other 3 individuals. Then $p(1) = P(Y = 1) = P(A \text{ first}) = \frac{1}{4} = .25$

$$p(2) = P(Y = 2) = P(B, C, \text{ or } D \text{ first and } A \text{ next}) = \frac{3}{4} \cdot \frac{1}{3} = \frac{1}{4} = .25$$

$$p(4) = P(Y = 3) = P(A \text{ last}) = \frac{3}{4} \cdot \frac{2}{3} \cdot \frac{1}{2} = \frac{1}{4} = .25$$

$$\text{So } p(3) = 1 - (.25 + .25 + .25) = .25$$

20. $P(0) = P(Y = 0) = P(\text{both arrive on Wed.}) = (.3)(.3) = .09$

$$P(1) = P(Y = 1) = P[(W, Th) \text{ or } (Th, W) \text{ or } (Th, Th)] \\ = (.3)(.4) + (.4)(.3) + (.4)(.4) = .40$$

$$P(2) = P(Y = 2) = P[(W, F) \text{ or } (Th, F) \text{ or } (F, W) \text{ or } (F, Th) \text{ or } (F, F)] = .32$$

$$P(3) = 1 - [.09 + .40 + .32] = .19$$

Chapter 3: Discrete Random Variables and Probability Distributions

21. The jumps in $F(x)$ occur at $x = 0, 1, 2, 3, 4, 5$, and 6 , so we first calculate $F(\cdot)$ at each of these values:

$$F(0) = P(X \leq 0) = P(X = 0) = .10$$

$$F(1) = P(X \leq 1) = p(0) + p(1) = .25$$

$$F(2) = P(X \leq 2) = p(0) + p(1) + p(2) = .45$$

$$F(3) = .70, F(4) = .90, F(5) = .96, \text{ and } F(6) = 1.$$

The c.d.f. is

$$F(x) = \begin{cases} .00 & x < 0 \\ .10 & 0 \leq x < 1 \\ .25 & 1 \leq x < 2 \\ .45 & 2 \leq x < 3 \\ .70 & 3 \leq x < 4 \\ .90 & 4 \leq x < 5 \\ .96 & 5 \leq x < 6 \\ 1.00 & 6 \leq x \end{cases}$$

Then $P(X \leq 3) = F(3) = .70$, $P(X < 3) = P(X \leq 2) = F(2) = .45$,

$P(3 \leq X) = 1 - P(X \leq 2) = 1 - F(2) = 1 - .45 = .55$,

and $P(2 \leq X \leq 5) = F(5) - F(1) = .96 - .25 = .71$

- 22.

a. $P(X = 2) = .39 - .19 = .20$

b. $P(X > 3) = 1 - .67 = .33$

c. $P(2 \leq X \leq 5) = .92 - .19 = .78$

d. $P(2 < X < 5) = .92 - .39 = .53$

- 23.

- a. Possible X values are those values at which $F(x)$ jumps, and the probability of any particular value is the size of the jump at that value. Thus we have:

x	1	3	4	6	12
$p(x)$	.30	.10	.05	.15	.40

b. $P(3 \leq X \leq 6) = F(6) - F(3-) = .60 - .30 = .30$

$$P(4 \leq X) = 1 - P(X < 4) = 1 - F(4-) = 1 - .40 = .60$$

- 24.

$$P(0) = P(Y = 0) = P(B \text{ first}) = p$$

$$P(1) = P(Y = 1) = P(G \text{ first, then } B) = P(GB) = (1 - p)p$$

$$P(2) = P(Y = 2) = P(GGB) = (1 - p)^2 p$$

Continuing, $p(y) = P(Y=y) = P(y \text{ G's and then a B}) = (1 - p)^y p$ for $y = 0, 1, 2, 3, \dots$

Chapter 3: Discrete Random Variables and Probability Distributions

25.

- a. Possible X values are 1, 2, 3, ...

$$P(1) = P(X = 1) = P(\text{return home after just one visit}) = \frac{1}{3}$$

$$P(2) = P(X = 2) = P(\text{second visit and then return home}) = \frac{2}{3} \cdot \frac{1}{3}$$

$$P(3) = P(X = 3) = P(\text{three visits and then return home}) = \left(\frac{2}{3}\right)^2 \cdot \frac{1}{3}$$

$$\text{In general } p(x) = \left(\frac{2}{3}\right)^{x-1} \left(\frac{1}{3}\right) \text{ for } x = 1, 2, 3, \dots$$

- b. The number of straight line segments is $Y = 1 + X$ (since the last segment traversed returns Alvie to O), so as in a, $p(y) = \left(\frac{2}{3}\right)^{y-2} \left(\frac{1}{3}\right)$ for $y = 2, 3, \dots$

- c. Possible Z values are 0, 1, 2, 3, ...

$$p(0) = P(\text{male first and then home}) = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6},$$

$$p(1) = P(\text{exactly one visit to a female}) = P(\text{female 1st, then home}) + P(F, M, \text{home}) + P(M, F, \text{home}) + P(M, F, M, \text{home})$$

$$= \left(\frac{1}{2}\right)\left(\frac{1}{3}\right) + \left(\frac{1}{2}\right)\left(\frac{2}{3}\right)\left(\frac{1}{3}\right) + \left(\frac{1}{2}\right)\left(\frac{2}{3}\right)\left(\frac{1}{3}\right) + \left(\frac{1}{2}\right)\left(\frac{2}{3}\right)\left(\frac{2}{3}\right)\left(\frac{1}{3}\right)$$

$$= \left(\frac{1}{2}\right)\left(1 + \frac{2}{3}\right)\left(\frac{1}{3}\right) + \left(\frac{1}{2}\right)\left(\frac{2}{3}\right)\left(\frac{2}{3} + 1\right)\left(\frac{1}{3}\right) = \left(\frac{1}{2}\right)\left(\frac{5}{3}\right)\left(\frac{1}{3}\right) + \left(\frac{1}{2}\right)\left(\frac{2}{3}\right)\left(\frac{5}{3}\right)\left(\frac{1}{3}\right)$$

where the first term corresponds to initially visiting a female and the second term corresponds to initially visiting a male. Similarly,

$$p(2) = \left(\frac{1}{2}\right)\left(\frac{2}{3}\right)^2 \left(\frac{5}{3}\right)\left(\frac{1}{3}\right) + \left(\frac{1}{2}\right)\left(\frac{2}{3}\right)^2 \left(\frac{5}{3}\right)\left(\frac{1}{3}\right). \text{ In general,}$$

$$p(z) = \left(\frac{1}{2}\right)\left(\frac{2}{3}\right)^{2z-2} \left(\frac{5}{3}\right)\left(\frac{1}{3}\right) + \left(\frac{1}{2}\right)\left(\frac{2}{3}\right)^{2z-2} \left(\frac{5}{3}\right)\left(\frac{1}{3}\right) = \left(\frac{24}{54}\right)\left(\frac{2}{3}\right)^{2z-2} \text{ for } z = 1, 2, 3, \dots$$

26.

- a. The sample space consists of all possible permutations of the four numbers 1, 2, 3, 4:

outcome	y value	outcome	y value	outcome	y value
1234	4	2314	1	3412	0
1243	2	2341	0	3421	0
1324	2	2413	0	4132	1
1342	1	2431	1	4123	0
1423	1	3124	1	4213	1
1432	2	3142	0	4231	2
2134	2	3214	2	4312	0
2143	0	3241	1	4321	0

- b. Thus $p(0) = P(Y = 0) = \frac{9}{24}$, $p(1) = P(Y = 1) = \frac{8}{24}$, $p(2) = P(Y = 2) = \frac{6}{24}$,
 $p(3) = P(Y = 3) = 0$, $p(4) = P(Y = 4) = \frac{1}{24}$.

Chapter 3: Discrete Random Variables and Probability Distributions

27. If $x_1 < x_2$, $F(x_2) = P(X \leq x_2) = P(\{X \leq x_1\} \cup \{x_1 < X \leq x_2\})$
 $= P(X \leq x_1) + P(x_1 < X \leq x_2) \geq P(X \leq x_1) = F(x_1)$.
 $F(x_1) = F(x_2)$ when $P(x_1 < X \leq x_2) = 0$.

Section 3.3

28.

a. $E(X) = \sum_{x=0}^4 x \cdot p(x)$
 $= (0)(.08) + (1)(.15) + (2)(.45) + (3)(.27) + (4)(.05) = 2.06$

b. $V(X) = \sum_{x=0}^4 (x - 2.06)^2 \cdot p(x) = (0 - 2.06)^2(.08) + \dots + (4 - 2.06)^2(.05)$
 $= .339488 + .168540 + .001620 + .238572 + .188180 = .9364$

c. $\sigma_x = \sqrt{.9364} = .9677$

d. $V(X) = \left[\sum_{x=0}^4 x^2 \cdot p(x) \right] - (2.06)^2 = 5.1800 - 4.2436 = .9364$

29.

a. $E(Y) = \sum_{y=0}^4 y \cdot p(y) = (0)(.60) + (1)(.25) + (2)(.10) + (3)(.05) = .60$

b. $E(100Y^2) = \sum_{y=0}^4 100y^2 \cdot p(y) = (0)(.60) + (100)(.25)$
 $+ (400)(.10) + (900)(.05) = 110$

30.

$E(Y) = .60$;
 $E(Y^2) = 1.1$
 $V(Y) = E(Y^2) - [E(Y)]^2 = 1.1 - (.60)^2 = .74$
 $\sigma_y = \sqrt{.74} = .8602$
 $E(Y) \pm \sigma_y = .60 \pm .8602 = (-.2602, 1.4602)$ or $(0, 1)$.
 $P(Y = 0) + P(Y = 1) = .85$

Chapter 3: Discrete Random Variables and Probability Distributions

31.

- a. $E(X) = (13.5)(.2) + (15.9)(.5) + (19.1)(.3) = 16.38$,
 $E(X^2) = (13.5)^2(.2) + (15.9)^2(.5) + (19.1)^2(.3) = 272.298$,
 $V(X) = 272.298 - (16.38)^2 = 3.9936$
- b. $E(25X - 8.5) = 25 E(X) - 8.5 = (25)(16.38) - 8.5 = 401$
- c. $V(25X - 8.5) = V(25X) = (25)^2 V(X) = (625)(3.9936) = 2496$
- d. $E[h(X)] = E[X - .01X^2] = E(X) - .01E(X^2) = 16.38 - 2.72 = 13.66$

32.

- a. $E(X^2) = \sum_{x=0}^1 x^2 \cdot p(x) = (0^2)((1-p) + (1^2)(p) = (1)(p) = p$
- b. $V(X) = E(X^2) - [E(X)]^2 = p - p^2 = p(1-p)$
- c. $E(x^{79}) = (0^{79})(1-p) + (1^{79})(p) = p$

33.

$E(X) = \sum_{x=1}^{\infty} x \cdot p(x) = \sum_{x=1}^{\infty} x \cdot \frac{c}{x^3} = c \sum_{x=1}^{\infty} \frac{1}{x^2}$, but it is a well-known result from the theory of infinite series that $\sum_{x=1}^{\infty} \frac{1}{x^2} < \infty$, so $E(X)$ is finite.

34.

Let $h(X)$ denote the net revenue (sales revenue – order cost) as a function of X . Then $h_3(X)$ and $h_4(X)$ are the net revenue for 3 and 4 copies purchased, respectively. For $x = 1$ or 2 , $h_3(X) = 2x - 3$, but at $x = 3, 4, 5, 6$ the revenue plateaus. Following similar reasoning, $h_4(X) = 2x - 4$ for $x = 1, 2, 3$, but plateaus at 4 for $x = 4, 5, 6$.

x	1	2	3	4	5	6
$h_3(x)$	-1	1	3	3	3	3
$h_4(x)$	-2	0	2	4	4	4
$p(x)$	$\frac{1}{15}$	$\frac{2}{15}$	$\frac{3}{15}$	$\frac{4}{15}$	$\frac{3}{15}$	$\frac{2}{15}$

$$E[h_3(X)] = \sum_{x=1}^6 h_3(x) \cdot p(x) = (-1)\left(\frac{1}{15}\right) + \dots + (3)\left(\frac{2}{15}\right) = 2.4667$$

$$\text{Similarly, } E[h_4(X)] = \sum_{x=1}^6 h_4(x) \cdot p(x) = (-2)\left(\frac{1}{15}\right) + \dots + (4)\left(\frac{2}{15}\right) = 2.6667$$

Ordering 4 copies gives slightly higher revenue, on the average.

Chapter 3: Discrete Random Variables and Probability Distributions

35.

P(x)	.8	.1	.08	.02
x	0	1,000	5,000	10,000
H(x)	0	500	4,500	9,500

$E[h(X)] = 600$. Premium should be \$100 plus expected value of damage minus deductible or \$700.

$$\begin{aligned}
 36. \quad E(X) &= \sum_{x=1}^n x \cdot \left(\frac{1}{n}\right) = \left(\frac{1}{n}\right) \sum_{x=1}^n x = \frac{1}{n} \left[\frac{n(n+1)}{2} \right] = \frac{n+1}{2} \\
 E(X^2) &= \sum_{x=1}^n x^2 \cdot \left(\frac{1}{n}\right) = \left(\frac{1}{n}\right) \sum_{x=1}^n x^2 = \frac{1}{n} \left[\frac{n(n+1)(2n+1)}{6} \right] = \frac{(n+1)(2n+1)}{6} \\
 \text{So } V(X) &= \frac{(n+1)(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2 = \frac{n^2-1}{12}
 \end{aligned}$$

$$37. \quad E[h(X)] = E\left(\frac{1}{X}\right) = \sum_{x=1}^6 \left(\frac{1}{x}\right) \cdot p(x) = \frac{1}{6} \sum_{x=1}^6 \frac{1}{x} = .408, \text{ whereas } \frac{1}{3.5} = .286, \text{ so you}$$

expect to win more if you gamble.

$$38. \quad E(X) = \sum_{x=1}^4 x \cdot p(x) = 2.3, E(X^2) = 6.1, \text{ so } V(X) = 6.1 - (2.3)^2 = .81$$

Each lot weighs 5 lbs, so weight left = $100 - 5x$.
 Thus the expected weight left is $100 - 5E(X) = 88.5$,
 and the variance of the weight left is
 $V(100 - 5X) = V(-5X) = 25V(x) = 20.25$.

39.

a. The line graph of the p.m.f. of $-X$ is just the line graph of the p.m.f. of X reflected about zero, but both have the same degree of spread about their respective means, suggesting $V(-X) = V(X)$.

b. With $a = -1$, $b = 0$, $V(aX + b) = V(-X) = a^2V(X)$.

$$\begin{aligned}
 40. \quad V(aX + b) &= \sum_x [aX + b - E(aX + b)]^2 \cdot p(x) = \sum_x [aX + b - (a\mu + b)]^2 p(x) \\
 &= \sum_x [aX - (a\mu)]^2 p(x) = a^2 \sum_x [X - \mu]^2 p(x) = a^2V(X).
 \end{aligned}$$

Chapter 3: Discrete Random Variables and Probability Distributions

41.

a. $E[X(X-1)] = E(X^2) - E(X), \quad \Rightarrow E(X^2) = E[X(X-1)] + E(X) = 32.5$

b. $V(X) = 32.5 - (5)^2 = 7.5$

c. $V(X) = E[X(X-1)] + E(X) - [E(X)]^2$

42. With $a = 1$ and $b = c$, $E(X - c) = E(aX + b) = aE(X) + b = E(X) - c$. When $c = \mu$, $E(X - \mu) = E(X) - \mu = \mu - \mu = 0$, so the expected deviation from the mean is zero.

43.

a.

k	2	3	4	5	10
$\frac{1}{k^2}$	.25	.11	.06	.04	.01

b. $\mu = \sum_{x=0}^6 x \cdot p(x) = 2.64, \quad \sigma^2 = \left[\sum_{x=0}^6 x^2 \cdot p(x) \right] - \mu^2 = 2.37, \quad \sigma = 1.54$

Thus $\mu - 2\sigma = -.44$, and $\mu + 2\sigma = 5.72$,

so $P(|x - \mu| \geq 2\sigma) = P(X \text{ is at least 2 s.d.'s from } \mu)$

$$= P(x \text{ is either } \leq -.44 \text{ or } \geq 5.72) = P(X = 6) = .04.$$

Chebyshev's bound of .025 is much too conservative. For $K = 3, 4, 5$, and 10, $P(|x - \mu| \geq k\sigma) = 0$, here again pointing to the very conservative nature of the bound $\frac{1}{k^2}$.

c. $\mu = 0$ and $\sigma = \frac{1}{3}$, so $P(|x - \mu| \geq 3\sigma) = P(|X| \geq 1)$

$$= P(X = -1 \text{ or } +1) = \frac{1}{18} + \frac{1}{18} = \frac{1}{9}, \text{ identical to the upper bound.}$$

d. Let $p(-1) = \frac{1}{50}$, $p(+1) = \frac{1}{50}$, $p(0) = \frac{24}{25}$.

Section 3.4

44.

$$\text{a. } b(3;8,.6) = \binom{8}{3} (.6)^3 (.4)^5 = (56)(.00221184) = .124$$

$$\text{b. } b(5;8,.6) = \binom{8}{5} (.6)^5 (.4)^3 = (56)(.00497664) = .279$$

$$\text{c. } P(3 \leq X \leq 5) = b(3;8,.6) + b(4;8,.6) + b(5;8,.6) = .635$$

$$\text{d. } P(1 \leq X) = 1 - P(X = 0) = 1 - \binom{12}{0} (.1)^0 (.9)^{12} = 1 - (.9)^{12} = .718$$

45.

$$\text{a. } B(4;10,.3) = .850$$

$$\text{b. } b(4;10,.3) = B(4;10,.3) - B(3;10,.3) = .200$$

$$\text{c. } b(6;10,.7) = B(6;10,.7) - B(5;10,.7) = .200$$

$$\text{d. } P(2 \leq X \leq 4) = B(4;10,.3) - B(1;10,.3) = .701$$

$$\text{e. } P(2 < X) = 1 - P(X \leq 1) = 1 - B(1;10,.3) = .851$$

$$\text{f. } P(X \leq 1) = B(1;10,.7) = .0000$$

$$\text{g. } P(2 < X < 6) = P(3 \leq X \leq 5) = B(5;10,.3) - B(2;10,.3) = .570$$

46. $X \sim \text{Bin}(25, .05)$

$$\text{a. } P(X \leq 2) = B(2;25,.05) = .873$$

$$\text{b. } P(X \geq 5) = 1 - P(X \leq 4) = 1 - B(4;25,.05) = .1 - .993 = .007$$

$$\text{c. } P(1 \leq X \leq 4) = P(X \leq 4) - P(X \leq 0) = .993 - .277 = .716$$

$$\text{d. } P(X = 0) = P(X \leq 0) = .277$$

$$\begin{aligned} \text{e. } E(X) &= np = (25)(.05) = 1.25 \\ V(X) &= np(1-p) = (25)(.05)(.95) = 1.1875 \\ \sigma_x &= 1.0897 \end{aligned}$$

Chapter 3: Discrete Random Variables and Probability Distributions

47. $X \sim \text{Bin}(6, .10)$

a. $P(X = 1) = \binom{n}{x} (p)^x (1 - p)^{n-x} = \binom{6}{1} (.1)^1 (.9)^5 = .3543$

b. $P(X \geq 2) = 1 - [P(X = 0) + P(X = 1)]$.

From a, we know $P(X = 1) = .3543$, and $P(X = 0) = \binom{6}{0} (.1)^0 (.9)^6 = .5314$.

Hence $P(X \geq 2) = 1 - [.3543 + .5314] = .1143$

c. Either 4 or 5 goblets must be selected

i) Select 4 goblets with zero defects: $P(X = 0) = \binom{4}{0} (.1)^0 (.9)^4 = .6561$.

ii) Select 4 goblets, one of which has a defect, and the 5th is good:

$$\left[\binom{4}{1} (.1)^1 (.9)^3 \right] \times .9 = .26244$$

So the desired probability is $.6561 + .26244 = .91854$

48. Let S = comes to a complete stop, so $p = .25$, $n = 20$

a. $P(X \leq 6) = B(6; 20, .25) = .786$

b. $P(X = 6) = b(6; 20, .25) = B(6; 20, .25) - B(5; 20, .25) = .786 - .617 = .169$

c. $P(X \geq 6) = 1 - P(X \leq 5) = 1 - B(5; 20, .25) = 1 - .617 = .383$

d. $E(X) = (20)(.25) = 5$. We expect 5 of the next 20 to stop.

49. Let S = has at least one citation. Then $p = .4$, $n = 15$

a. If at least 10 have no citations (Failure), then at most 5 have had at least one (Success):
 $P(X \leq 5) = B(5; 15, .40) = .403$

b. $P(X \leq 7) = B(7; 15, .40) = .787$

c. $P(5 \leq X \leq 10) = P(X \leq 10) - P(X \leq 4) = .991 - .217 = .774$

Chapter 3: Discrete Random Variables and Probability Distributions

50. $X \sim \text{Bin}(10, .60)$
- $P(X \geq 6) = 1 - P(X \leq 5) = 1 - B(5; 20, .60) = 1 - .367 = .633$
 - $E(X) = np = (10)(.6) = 6$; $V(X) = np(1 - p) = (10)(.6)(.4) = 2.4$;
 $\sigma_x = 1.55$
 $E(X) \pm \sigma_x = (4.45, 7.55)$.
 We desire $P(5 \leq X \leq 7) = P(X \leq 7) - P(X \leq 4) = .833 - .166 = .667$
 - $P(3 \leq X \leq 7) = P(X \leq 7) - P(X \leq 2) = .833 - .012 = .821$
51. Let S represent a telephone that is submitted for service while under warranty and must be replaced. Then $p = P(S) = P(\text{replaced} | \text{submitted}) \cdot P(\text{submitted}) = (.40)(.20) = .08$. Thus X, the number among the company's 10 phones that must be replaced, has a binomial distribution with $n = 10$, $p = .08$, so $p(2) = P(X=2) = \binom{10}{2} (.08)^2 (.92)^8 = .1478$
52. $X \sim \text{Bin}(25, .02)$
- $P(X=1) = 25(.02)(.98)^{24} = .308$
 - $P(X \geq 1) = 1 - P(X=0) = 1 - (.98)^{25} = 1 - .603 = .397$
 - $P(X \geq 2) = 1 - P(X \leq 1) = 1 - [.308 + .397]$
 - $\bar{x} = 25(.02) = .5$; $\sigma = \sqrt{npq} = \sqrt{25(.02)(.98)} = \sqrt{.49} = .7$
 $\bar{x} + 2\sigma = .5 + 1.4 = 1.9$ So $P(0 \leq X \leq 1.9) = P(X \leq 1) = .705$
 - $$\frac{.5(4.5) + 24.5(3)}{25} = 3.03 \text{ hours}$$
53. X = the number of flashlights that work.
 Let event $B = \{\text{battery has acceptable voltage}\}$.
 Then $P(\text{flashlight works}) = P(\text{both batteries work}) = P(B)P(B) = (.9)(.9) = .81$ We must assume that the batteries' voltage levels are independent.
 $X \sim \text{Bin}(10, .81)$. $P(X \geq 9) = P(X=9) + P(X=10)$

$$\binom{10}{9} (.81)^9 (.19) + \binom{10}{10} (.81)^{10} = .285 + .122 = .407$$

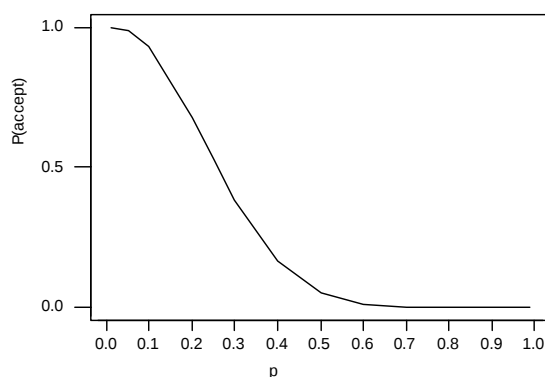
Chapter 3: Discrete Random Variables and Probability Distributions

54. Let p denote the actual proportion of defectives in the batch, and X denote the number of defectives in the sample.

a. $P(\text{the batch is accepted}) = P(X \leq 2) = B(2;10,p)$

p	.01	.05	.10	.20	.25
$P(\text{accept})$	1.00	.988	.930	.678	.526

b.



c. $P(\text{the batch is accepted}) = P(X \leq 1) = B(1;10,p)$

p	.01	.05	.10	.20	.25
$P(\text{accept})$	.996	.914	.736	.376	.244

d. $P(\text{the batch is accepted}) = P(X \leq 2) = B(2;15,p)$

p	.01	.05	.10	.20	.25
$P(\text{accept})$	1.00	.964	.816	.398	.236

- e. We want a plan for which $P(\text{accept})$ is high for $p \leq .1$ and low for $p > .1$. The plan in **d** seems most satisfactory in these respects.

Chapter 3: Discrete Random Variables and Probability Distributions

55.

- a. $P(\text{rejecting claim when } p = .8) = B(15; 25, .8) = .017$
- b. $P(\text{not rejecting claim when } p = .7) = P(X \geq 16 \text{ when } p = .7)$
 $= 1 - B(15; 25, .7) = 1 - .189 = .811$; for $p = .6$, this probability is
 $= 1 - B(15; 25, .6) = 1 - .575 = .425$.
- c. The probability of rejecting the claim when $p = .8$ becomes $B(14; 25, .8) = .006$, smaller than in **a** above. However, the probabilities of **b** above increase to .902 and .586, respectively.

56.

$$h(x) = 1 \cdot X + 2.25(25 - X) = 62.5 - 1.5X, \text{ so } E(h(X)) = 62.5 - 1.5E(x) \\ = 62.5 - 1.5np - 62.5 - (1.5)(25)(.6) = \$40.00$$

57.

If topic A is chosen, when $n = 2$, $P(\text{at least half received})$
 $= P(X \geq 1) = 1 - P(X = 0) = 1 - (.1)^2 = .99$
 If B is chosen, when $n = 4$, $P(\text{at least half received})$
 $= P(X \geq 2) = 1 - P(X \leq 1) = 1 - (0.1)^4 - 4(.1)^3(.9) = .9963$
 Thus topic B should be chosen.
 If $p = .5$, the probabilities are .75 for A and .6875 for B, so now A should be chosen.

58.

- a. $np(1 - p) = 0$ if either $p = 0$ (whence every trial is a failure, so there is no variability in X) or if $p = 1$ (whence every trial is a success and again there is no variability in X)
- b. $\frac{d}{dp}[np(1 - p)] = n[(1 - p) + p(-1)] = n[1 - 2p] = 0 \Rightarrow p = .5$, which is easily seen to correspond to a maximum value of $V(X)$.

59.

$$\text{a. } b(x; n, 1 - p) = \binom{n}{x} (1 - p)^x (p)^{n-x} = \binom{n}{n-x} (p)^{n-x} (1 - p)^x = b(n-x; n, p)$$

Alternatively, $P(x \text{ S's when } P(S) = 1 - p) = P(n-x \text{ F's when } P(F) = p)$, since the two events are identical, but the labels S and F are arbitrary so can be interchanged (if $P(S)$ and $P(F)$ are also interchanged), yielding $P(n-x \text{ S's when } P(S) = 1 - p)$ as desired.

- b. $B(x; n, 1 - p) = P(\text{at most } x \text{ S's when } P(S) = 1 - p)$
 $= P(\text{at least } n-x \text{ F's when } P(F) = p)$
 $= P(\text{at least } n-x \text{ S's when } P(S) = p)$
 $= 1 - P(\text{at most } n-x-1 \text{ S's when } P(S) = p)$
 $= 1 - B(n-x-1; n, p)$
- c. Whenever $p > .5$, $(1 - p) < .5$ so probabilities involving X can be calculated using the results **a** and **b** in combination with tables giving probabilities only for $p \leq .5$

Chapter 3: Discrete Random Variables and Probability Distributions

60. Proof of $E(X) = np$:

$$\begin{aligned}
 E(X) &= \sum_{x=0}^n x \cdot \binom{n}{x} p^x (1-p)^{n-x} = \sum_{x=1}^n x \cdot \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \\
 &= \sum_{x=1}^n \frac{n!}{(x-1)!(n-x)!} p^x (1-p)^{n-x} = np \sum_{x=1}^n \frac{(n-1)!}{(x-1)!(n-x)!} p^{x-1} (1-p)^{n-x} \\
 &= np \sum_{y=0}^{n-1} \frac{(n-1)!}{(y)!(n-1-y)!} p^y (1-p)^{n-1-y} \quad (y \text{ replaces } x-1) \\
 &= np \left\{ \sum_{y=0}^{n-1} \binom{n-1}{y} p^y (1-p)^{n-1-y} \right\}
 \end{aligned}$$

The expression in braces is the sum over all possible values $y = 0, 1, 2, \dots, n-1$ of a binomial p.m.f. based on $n-1$ trials, so equals 1, leaving only np , as desired.

61.

- a. Although there are three payment methods, we are only concerned with S = uses a debit card and F = does not use a debit card. Thus we can use the binomial distribution. So $n = 100$ and $p = .5$. $E(X) = np = 100(.5) = 50$, and $V(X) = 25$.
- b. With S = doesn't pay with cash, $n = 100$ and $p = .7$, $E(X) = np = 100(.7) = 70$, and $V(X) = 21$.

62.

- a. Let X = the number with reservations who show, a binomial r.v. with $n = 6$ and $p = .8$.
The desired probability is
 $P(X = 5 \text{ or } 6) = b(5;6,.8) + b(6;6,.8) = .3932 + .2621 = .6553$
- b. Let $h(X)$ = the number of available spaces. Then

When x is:	0	1	2	3	4	5	6
$H(x)$ is:	4	3	2	1	0	0	0

$$E[h(X)] = \sum_{x=0}^6 h(x) \cdot b(x;6,.8) = 4(.000) + 3(.002) = 2(.015 + 3(.082) = .277$$
- c. Possible X values are 0, 1, 2, 3, and 4. $X = 0$ if there are 3 reservations and none show or ...or 6 reservations and none show, so
 $P(X = 0) = b(0;3,.8)(.1) + b(0;4,.8)(.2) + b(0;5,.8)(.3) + b(0;6,.8)(.4)$
 $= .0080(.1) + .0016(.2) + .0003(.3) + .0001(.4) = .0013$
 $P(X = 1) = b(1;3,.8)(.1) + \dots + b(1;6,.8)(.4) = .0172$
 $P(X = 2) = .0906, \quad P(X = 3) = .2273,$
 $P(X = 4) = 1 - [.0013 + \dots + .2273] = .6636$

Chapter 3: Discrete Random Variables and Probability Distributions

63. When $p = .5$, $\mu = 10$ and $\sigma = 2.236$, so $2\sigma = 4.472$ and $3\sigma = 6.708$.
The inequality $|X - 10| \geq 4.472$ is satisfied if either $X \leq 5$ or $X \geq 15$, or $P(|X - \mu| \geq 2\sigma) = P(X \leq 5 \text{ or } X \geq 15) = .021 + .021 = .042$.

In the case $p = .75$, $\mu = 15$ and $\sigma = 1.937$, so $2\sigma = 3.874$ and $3\sigma = 5.811$. $P(|X - 15| \geq 3.874) = P(X \leq 11 \text{ or } X \geq 19) = .041 + .024 = .065$, whereas $P(|X - 15| \geq 5.811) = P(X \leq 9) = .004$. All these probabilities are considerably less than the upper bounds .25 (for $k = 2$) and .11 (for $k = 3$) given by Chebyshev.

Section 3.5

64.

- a. $X \sim \text{Hypergeometric } N=15, n=5, M=6$

$$\text{b. } P(X=2) = \frac{\binom{6}{2}\binom{9}{3}}{\binom{15}{5}} = \frac{840}{3003} = .280$$

$$P(X \leq 2) = P(X=0) + P(X=1) + P(X=2)$$

$$= \frac{\binom{9}{5}}{\binom{15}{5}} + \frac{\binom{6}{1}\binom{9}{4}}{\binom{15}{5}} + \frac{840}{3003} = \frac{126 + 756 + 840}{3003} = \frac{1722}{3003} = .573$$

$$P(X \geq 2) = 1 - P(X \leq 1) = 1 - [P(X=0) + P(X=1)] = 1 - \frac{126 + 756}{3003} = .706$$

$$\text{c. } E(X) = 5 \left(\frac{6}{15} \right) = 2; V(X) = \left(\frac{15-5}{14} \right) \cdot 5 \cdot \left(\frac{6}{15} \right) \cdot \left(1 - \frac{6}{15} \right) = .857;$$

$$\sigma = \sqrt{V(X)} = .926$$

Chapter 3: Discrete Random Variables and Probability Distributions

65. $X \sim h(x; 6, 12, 7)$

$$\text{a. } P(X=5) = \frac{\binom{7}{5} \binom{5}{1}}{\binom{12}{6}} = \frac{105}{924} = .114$$

$$\text{b. } P(X \leq 4) = 1 - P(X \geq 5) = 1 - [P(X=5) + P(X=6)] =$$

$$1 - \left[\frac{\binom{7}{5} \binom{5}{1}}{\binom{12}{6}} + \frac{\binom{7}{6} \binom{5}{0}}{\binom{12}{6}} \right] = 1 - \frac{105 + 7}{924} = 1 - .121 = .879$$

$$\text{c. } E(X) = \left(\frac{6 \cdot 7}{12} \right) = 3.5; \sigma = \sqrt{\left(\frac{6}{11} \right) \left(6 \right) \left(\frac{7}{12} \right) \left(\frac{5}{12} \right)} = \sqrt{.795} = .892$$

$$P(X > 3.5 + .892) = P(X > 4.392) = P(X \geq 5) = .121 \text{ (see part b)}$$

d. We can approximate the hypergeometric distribution with the binomial if the population size and the number of successes are large: $h(x; 15, 40, 400)$ approaches $b(x; 15, .10)$. So $P(X \leq 5) \approx B(5; 15, .10)$ from the binomial tables = .998

66.

$$\text{a. } P(X = 10) = h(10; 15, 30, 50) = \frac{\binom{30}{10} \binom{20}{5}}{\binom{50}{15}} = .2070$$

$$\text{b. } P(X \geq 10) = h(10; 15, 30, 50) + h(11; 15, 30, 50) + \dots + h(15; 15, 30, 50)$$

$$= .2070 + .1176 + .0438 + .0101 + .0013 + .0001 = .3799$$

c. $P(\text{at least 10 from the same class}) = P(\text{at least 10 from second class [answer from b]}) + P(\text{at least 10 from first class})$. But “at least 10 from 1st class” is the same as “at most 5 from the second” or $P(X \leq 5)$.

$$P(X \leq 5) = h(0; 15, 30, 50) + h(1; 15, 30, 50) + \dots + h(5; 15, 30, 50)$$

$$= .11697 + .002045 + .000227 + .000150 + .000001 + .000000$$

$$= .01412$$

$$\text{So the desired probability} = P(x \geq 10) + P(X \leq 5)$$

$$= .3799 + .01412 = .39402$$

Chapter 3: Discrete Random Variables and Probability Distributions

d. $E(X) = n \cdot \frac{M}{N} = 15 \cdot \frac{30}{50} = 9$

$$V(X) = \left(\frac{35}{49}\right)(9)\left(1 - \frac{30}{50}\right) = 2.5714$$

$$\sigma_x = 1.6036$$

- e. Let $Y = 15 - X$. Then $E(Y) = 15 - E(X) = 15 - 9 = 6$
 $V(Y) = V(15 - X) = V(X) = 2.5714$, so $\sigma_Y = 1.6036$

67.

- a. Possible values of X are 5, 6, 7, 8, 9, 10. (In order to have less than 5 of the granite, there would have to be more than 10 of the basaltic).

$$P(X = 5) = h(5; 15, 10, 20) = \frac{\binom{10}{5} \binom{10}{10}}{\binom{20}{15}} = .0163.$$

Following the same pattern for the other values, we arrive at the pmf, in table form below.

x	5	6	7	8	9	10
p(x)	.0163	.1354	.3483	.3483	.1354	.0163

- b. $P(\text{all 10 of one kind or the other}) = P(X = 5) + P(X = 10) = .0163 + .0163 = .0326$

c. $E(X) = n \cdot \frac{M}{N} = 15 \cdot \frac{10}{20} = 7.5$; $V(X) = \left(\frac{5}{19}\right)(7.5)\left(1 - \frac{10}{20}\right) = .9868$;

$$\sigma_x = .9934$$

$$\mu \pm \sigma = 7.5 \pm .9934 = (6.5066, 8.4934), \text{ so we want } P(X = 7) + P(X = 8) = .3483 + .3483 = .6966$$

68.

- a. $h(x; 6, 4, 11)$

b. $6 \cdot \left(\frac{4}{11}\right) = 2.18$

Chapter 3: Discrete Random Variables and Probability Distributions

69.

- a. $h(x; 10, 10, 20)$ (the successes here are the top 10 pairs, and a sample of 10 pairs is drawn from among the 20)
- b. Let X = the number among the top 5 who play E-W. Then $P(\text{all of top 5 play the same direction}) = P(X = 5) + P(X = 0) = h(5; 10, 5, 20) + h(5; 10, 5, 20)$

$$= \frac{\binom{15}{5}}{\binom{20}{10}} + \frac{\binom{15}{10}}{\binom{20}{10}} = .033$$

- c. $N = 2n$; $M = n$; $n = n$
 $h(x; n, n, 2n)$

$$E(X) = n \cdot \frac{n}{2n} = \frac{1}{2}n;$$

$$V(X) =$$

$$\left(\frac{2n-n}{2n-1}\right) \cdot n \cdot \frac{n}{2n} \cdot \left(1 - \frac{n}{2n}\right) = \left(\frac{n}{2n-1}\right) \cdot \frac{n}{2} \cdot \left(1 - \frac{n}{2n}\right) = \left(\frac{n}{2n-1}\right) \cdot \frac{n}{2} \cdot \left(\frac{1}{2}\right)$$

70.

- a. $h(x; 10, 15, 50)$

- b. When N is large relative to n , $h(x; n, M, N) \doteq b\left(x; n, \frac{M}{N}\right)$,

$$\text{so } h(x; 10, 150, 500) \doteq b(x; 10, .3)$$

- c. Using the hypergeometric model, $E(X) = 10 \cdot \left(\frac{150}{500}\right) = 3$ and

$$V(X) = \frac{490}{499}(10)(.3)(.7) = .982(2.1) = 2.06$$

$$\text{Using the binomial model, } E(X) = (10)(.3) = 3, \text{ and } V(X) = 10(.3)(.7) = 2.1$$

Chapter 3: Discrete Random Variables and Probability Distributions

71.

a. With S = a female child and F = a male child, let X = the number of F's before the 2nd S.
Then $P(X = x) = nb(x; 2, .5)$

b. $P(\text{exactly 4 children}) = P(\text{exactly 2 males})$
 $= nb(2; 2, .5) = (3)(.0625) = .188$

c. $P(\text{at most 4 children}) = P(X \leq 2)$
 $= \sum_{x=0}^2 nb(x; 2, .5) = .25 + 2(.25)(.5) + 3(.0625) = .688$

d. $E(X) = \frac{(2)(.5)}{.5} = 2$, so the expected number of children = $E(X + 2)$
 $= E(X) + 2 = 4$

72.

The only possible values of X are 3, 4, and 5.

$p(3) = P(X = 3) = P(\text{first 3 are B's or first 3 are G's}) = 2(.5)^3 = .250$

$p(4) = P(\text{two among the 1st three are B's and the 4th is a B}) + P(\text{two among the 1st three are$

G's and the 4th is a G) = $2 \cdot \binom{3}{2} (.5)^4 = .375$

$p(5) = 1 - p(3) - p(4) = .375$

73.

This is identical to an experiment in which a single family has children until exactly 6 females have been born (since $p = .5$ for each of the three families), so $p(x) = nb(x; 6, .5)$ and $E(X) = 6$ ($= 2+2+2$, the sum of the expected number of males born to each one.)

74.

The interpretation of "roll" here is a pair of tosses of a single player's die (two tosses by A or two by B). With S = doubles on a particular roll, $p = \frac{1}{6}$. Furthermore, A and B are really identical (each die is fair), so we can equivalently imagine A rolling until 10 doubles appear. The $P(x \text{ rolls}) = P(9 \text{ doubles among the first } x - 1 \text{ rolls and a double on the } x^{\text{th}} \text{ roll}) =$

$$\binom{x-1}{9} \left(\frac{5}{6}\right)^{x-10} \left(\frac{1}{6}\right)^9 \cdot \left(\frac{1}{6}\right) = \binom{x-1}{9} \left(\frac{5}{6}\right)^{x-10} \left(\frac{1}{6}\right)^{10}$$

$$E(X) = \frac{r(1-p)}{p} = \frac{10(\frac{5}{6})}{\frac{1}{6}} = 10(5) = 50 \quad V(X) = \frac{r(1-p)}{p^2} = \frac{10(\frac{5}{6})}{(\frac{1}{6})^2} = 10(5)(6) = 300$$

Section 3.6

75.

- a. $P(X \leq 8) = F(8;5) = .932$
- b. $P(X = 8) = F(8;5) - F(7;5) = .065$
- c. $P(X \geq 9) = 1 - P(X \leq 8) = .068$
- d. $P(5 \leq X \leq 8) = F(8;5) - F(4;5) = .492$
- e. $P(5 < X < 8) = F(7;5) - F(5;5) = .867 - .616 = .251$

76.

- a. $P(X \leq 5) = F(5;8) = .191$
- b. $P(6 \leq X \leq 9) = F(9;8) - F(5;8) = .526$
- c. $P(X \geq 10) = 1 - P(X \leq 9) = .283$
- d. $E(X) = \lambda = 10$, $\sigma_X = \sqrt{\lambda} = 2.83$, so $P(X > 12.83) = P(X \geq 13) = 1 - P(X \leq 12) = 1 - .936 = .064$

77.

- a. $P(X \leq 10) = F(10;20) = .011$
- b. $P(X > 20) = 1 - F(20;20) = 1 - .559 = .441$
- c. $P(10 \leq X \leq 20) = F(20;20) - F(9;20) = .559 - .005 = .554$
 $P(10 < X < 20) = F(19;20) - F(10;20) = .470 - .011 = .459$
- d. $E(X) = \lambda = 20$, $\sigma_X = \sqrt{\lambda} = 4.472$
 $P(\mu - 2\sigma < X < \mu + 2\sigma) = P(20 - 8.944 < X < 20 + 8.944)$
 $= P(11.056 < X < 28.944)$
 $= P(X \leq 28) - P(X \leq 11)$
 $= F(28;20) - F(11;20)$
 $= .966 - .021 = .945$

78.

- a. $P(X = 1) = F(1;2) - F(0;2) = .982 - .819 = .163$
- b. $P(X \geq 2) = 1 - P(X \leq 1) = 1 - F(1;2) = 1 - .982 = .018$
- c. $P(1^{\text{st}} \text{ doesn't} \cap 2^{\text{nd}} \text{ doesn't}) = P(1^{\text{st}} \text{ doesn't}) \cdot P(2^{\text{nd}} \text{ doesn't})$
 $= (.819)(.819) = .671$

Chapter 3: Discrete Random Variables and Probability Distributions

79. $p = \frac{1}{200}$; $n = 1000$; $\lambda = np = 5$
- $P(5 \leq X \leq 8) = F(8;5) - F(4;5) = .492$
 - $P(X \geq 8) = 1 - P(X \leq 7) = 1 - .867 = .133$
- 80.
- The experiment is binomial with $n = 10,000$ and $p = .001$,
so $\mu = np = 10$ and $\sigma = \sqrt{npq} = 3.161$.
 - X has approximately a Poisson distribution with $\lambda = 10$,
so $P(X > 10) \approx 1 - F(10;10) = 1 - .583 = .417$
 - $P(X = 0) \approx 0$
- 81.
- $\lambda = 8$ when $t = 1$, so $P(X = 6) = F(6;8) - F(5;8) = .313 - .191 = .122$,
 $P(X \geq 6) = 1 - F(5;8) = .809$, and $P(X \geq 10) = 1 - F(9;8) = .283$
 - $t = 90 \text{ min} = 1.5 \text{ hours}$, so $\lambda = 12$; thus the expected number of arrivals is 12 and the SD
 $= \sqrt{12} = 3.464$
 - $t = 2.5 \text{ hours}$ implies that $\lambda = 20$; in this case, $P(X \geq 20) = 1 - F(19;20) = .530$ and $P(X \leq 10) = F(10;20) = .011$.
- 82.
- $P(X = 4) = F(4;5) - F(3;5) = .440 - .265 = .175$
 - $P(X \geq 4) = 1 - P(X \leq 3) = 1 - .265 = .735$
 - Arrivals occur at the rate of 5 per hour, so for a 45 minute period the rate is $\lambda = (5)(.75) = 3.75$, which is also the expected number of arrivals in a 45 minute period.
- 83.
- For a two hour period the parameter of the distribution is $\lambda t = (4)(2) = 8$,
so $P(X = 10) = F(10;8) - F(9;8) = .099$.
 - For a 30 minute period, $\lambda t = (4)(.5) = 2$, so $P(X = 0) = F(0,2) = .135$
 - $E(X) = \lambda t = 2$

Chapter 3: Discrete Random Variables and Probability Distributions

- 84.** Let X = the number of diodes on a board that fail.
- $E(X) = np = (200)(.01) = 2$, $V(X) = npq = (200)(.01)(.99) = 1.98$, $\sigma_X = 1.407$
 - X has approximately a Poisson distribution with $\lambda = np = 2$,
so $P(X \geq 4) = 1 - P(X \leq 3) = 1 - F(3;2) = 1 - .857 = .143$
 - $P(\text{board works properly}) = P(\text{all diodes work}) = P(X = 0) = F(0;2) = .135$
Let Y = the number among the five boards that work, a binomial r.v. with $n = 5$ and $p = .135$. Then $P(Y \geq 4) = P(Y = 4) + P(Y = 5) =$

$$\binom{5}{4}(.135)^4(.865) + \binom{5}{5}(.135)^5(.865)^0 = .00144 + .00004 = .00148$$
- 85.** $\alpha = 1/(\text{mean time between occurrences}) = \frac{1}{.5} = 2$
- $\alpha t = (2)(2) = 4$
 - $P(X > 5) = 1 - P(X \leq 5) = 1 - .785 = .215$
 - Solve for t , given $\alpha = 2$:
 $.1 = e^{-\alpha t}$
 $\ln(.1) = -\alpha t$
 $t = \frac{2.3026}{2} \approx 1.15 \text{ years}$
- 86.**
$$E(X) = \sum_{x=0}^{\infty} x \frac{e^{-\lambda} \lambda^x}{x!} = \sum_{x=1}^{\infty} x \frac{e^{-\lambda} \lambda^x}{x!} = \lambda \sum_{x=1}^{\infty} x \frac{e^{-\lambda} \lambda^{x-1}}{x!} = \lambda \sum_{y=0}^{\infty} y \frac{e^{-\lambda} \lambda^y}{y!} = \lambda$$
- 87.**
- For a one-quarter acre plot, the parameter is $(80)(.25) = 20$,
so $P(X \leq 16) = F(16;20) = .221$
 - The expected number of trees is $\lambda \cdot (\text{area}) = 80(85,000) = 6,800,000$.
 - The area of the circle is $\pi r^2 = .031416$ sq. miles or 20.106 acres. Thus X has a Poisson distribution with parameter 20.106

88.

$$\begin{aligned} \text{a. } P(X = 10 \text{ and no violations}) &= P(\text{no violations} \mid X = 10) \cdot P(X = 10) \\ &= (.5)^{10} \cdot [F(10;10) - F(9;10)] \\ &= (.000977)(.125) = .000122 \end{aligned}$$

$$\begin{aligned} \text{b. } P(y \text{ arrive and exactly 10 have no violations}) &= P(\text{exactly 10 have no violations} \mid y \text{ arrive}) \cdot P(y \text{ arrive}) \\ &= P(10 \text{ successes in } y \text{ trials when } p = .5) \cdot e^{-10} \frac{(10)^y}{y!} \\ &= \binom{y}{10} (.5)^{10} (.5)^{y-10} e^{-10} \frac{(10)^y}{y!} = \frac{e^{-10} (5)^y}{10!(y-10)!} \end{aligned}$$

$$\begin{aligned} \text{c. } P(\text{exactly 10 without a violation}) &= \sum_{y=10}^{\infty} \frac{e^{-10} (5)^y}{10!(y-10)!} \\ &= \frac{e^{-10} \cdot 5^{10}}{10!} \sum_{y=10}^{\infty} \frac{(5)^{y-10}}{(y-10)!} = \frac{e^{-10} \cdot 5^{10}}{10!} \sum_{u=0}^{\infty} \frac{(5)^u}{(u)!} = \frac{e^{-10} \cdot 5^{10}}{10!} \cdot e^5 \\ &= \frac{e^{-5} \cdot 5^{10}}{10!} = p(10;5). \end{aligned}$$

In fact, generalizing this argument shows that the number of “no-violation” arrivals within the hour has a Poisson distribution with parameter 5; the 5 results from $\lambda p = 10(.5)$.

89.

$$\begin{aligned} \text{a. } \text{No events in } (0, t+\Delta t) \text{ if and only if no events in } (0, t) \text{ and no events in } (t, t+\Delta t). \text{ Thus, } P_0(t+\Delta t) &= P_0(t) \cdot P(\text{no events in } (t, t+\Delta t)) \\ &= P_0(t)[1 - \lambda \cdot \Delta t - o(\Delta t)] \end{aligned}$$

$$\text{b. } \frac{P_0(t + \Delta t) - P_0(t)}{\Delta t} = -\lambda P_0(t) \frac{\Delta t}{\Delta t} - P_0(t) \cdot \frac{o(\Delta t)}{\Delta t}$$

$$\text{c. } \frac{d}{dt} [e^{-\lambda t}] = -\lambda e^{-\lambda t} = -\lambda P_0(t), \text{ as desired.}$$

$$\begin{aligned} \text{d. } \frac{d}{dt} \left[\frac{e^{-\lambda t} (\lambda t)^k}{k!} \right] &= \frac{-\lambda e^{-\lambda t} (\lambda t)^k}{k!} + \frac{k \lambda e^{-\lambda t} (\lambda t)^{k-1}}{k!} \\ &= -\lambda \frac{e^{-\lambda t} (\lambda t)^k}{k!} + \lambda \frac{e^{-\lambda t} (\lambda t)^{k-1}}{(k-1)!} = -\lambda P_k(t) + \lambda P_{k-1}(t) \text{ as desired.} \end{aligned}$$

Supplementary Exercises

90. Outcomes are (1,2,3)(1,2,4) (1,2,5) ... (5,6,7); there are 35 such outcomes. Each having probability $\frac{1}{35}$. The W values for these outcomes are 6 (=1+2+3), 7, 8, ..., 18. Since there is just one outcome with W value 6, $p(6) = P(W = 6) = \frac{1}{35}$. Similarly, there are three outcomes with W value 9 [(1,2,6) (1,3,5) and 2,3,4)], so $p(9) = \frac{3}{35}$. Continuing in this manner yields the following distribution:

W	6	7	8	9	10	11	12	13	14	15	16	17	18
P(W)	$\frac{1}{35}$	$\frac{1}{35}$	$\frac{2}{35}$	$\frac{3}{35}$	$\frac{4}{35}$	$\frac{4}{35}$	$\frac{5}{35}$	$\frac{4}{35}$	$\frac{4}{35}$	$\frac{3}{35}$	$\frac{2}{35}$	$\frac{1}{35}$	$\frac{1}{35}$

Since the distribution is symmetric about 12, $\mu = 12$, and $\sigma^2 = \sum_{w=6}^{18} (w-12)^2 p(w)$

$$= \frac{1}{35} [(6)^2(1) + (5)^2(1) + \dots + (5)^2(1) + (6)^2(1)] = 8$$

91.

- a. $p(1) = P(\text{exactly one suit}) = P(\text{all spades}) + P(\text{all hearts}) + P(\text{all diamonds})$

$$+ P(\text{all clubs}) = 4P(\text{all spades}) = 4 \cdot \frac{\binom{13}{5}}{\binom{52}{5}} = .00198$$

$p(2) = P(\text{all hearts and spades with at least one of each}) + \dots + P(\text{all diamonds and clubs with at least one of each})$

$= 6 P(\text{all hearts and spades with at least one of each})$

$= 6 [P(1 \text{ h and } 4 \text{ s}) + P(2 \text{ h and } 3 \text{ s}) + P(3 \text{ h and } 2 \text{ s}) + P(4 \text{ h and } 1 \text{ s})]$

$$= 6 \cdot \left[2 \cdot \frac{\binom{13}{4} \binom{13}{1}}{\binom{52}{5}} + 2 \cdot \frac{\binom{13}{3} \binom{13}{2}}{\binom{52}{5}} \right] = 6 \left[\frac{18,590 + 44,616}{2,598,960} \right] = .14592$$

$$p(4) = 4P(2 \text{ spades, } 1 \text{ h, } 1 \text{ d, } 1 \text{ c}) = \frac{4 \cdot \binom{13}{2} (13)(13)(13)}{\binom{52}{5}} = .26375$$

$$p(3) = 1 - [p(1) + p(2) + p(4)] = .58835$$

- b. $\mu = \sum_{x=1}^4 x \cdot p(x) = 3.114$, $\sigma^2 = \left[\sum_{x=1}^4 x^2 \cdot p(x) \right] - (3.114)^2 = .405$, $\sigma = .636$

Chapter 3: Discrete Random Variables and Probability Distributions

- 92.** $p(y) = P(Y = y) = P(y \text{ trials to achieve } r \text{ S's}) = P(y-r \text{ F's before } r^{\text{th}} \text{ S})$
 $= nb(y-r; r, p) = \binom{y-1}{r-1} p^r (1-p)^{y-r}, y = r, r+1, r+2, \dots$
- 93.**
- a.** $b(x; 15, .75)$
 - b.** $P(X > 10) = 1 - B(9; 15, .75) = 1 - .148$
 - c.** $B(10; 15, .75) - B(5; 15, .75) = .314 - .001 = .313$
 - d.** $\mu = (15)(.75) = 11.75, \sigma^2 = (15)(.75)(.25) = 2.81$
 - e.** Requests can all be met if and only if $X \leq 10$, and $15 - X \leq 8$, i.e. if $7 \leq X \leq 10$, so $P(\text{all requests met}) = B(10; 15, .75) - B(6; 15, .75) = .310$
- 94.** $P(6\text{-v light works}) = P(\text{at least one } 6\text{-v battery works}) = 1 - P(\text{neither works})$
 $= 1 - (1-p)^2$. $P(D \text{ light works}) = P(\text{at least } 2 \text{ d batteries work}) = 1 - P(\text{at most } 1 \text{ D battery works})$
 $= 1 - [(1-p)^4 + 4(1-p)^3]$. The 6-v should be taken if $1 - (1-p)^2 \geq 1 - [(1-p)^4 + 4(1-p)^3]$.
Simplifying, $1 \leq (1-p)^2 + 4p(1-p) \Rightarrow 0 \leq 2p - 3p^3 \Rightarrow p \leq \frac{2}{3}$.
- 95.** Let $X \sim \text{Bin}(5, .9)$. Then $P(X \geq 3) = 1 - P(X \leq 2) = 1 - B(2; 5, .9) = .991$
- 96.**
- a.** $P(X \geq 5) = 1 - B(4; 25, .05) = .007$
 - b.** $P(X \geq 5) = 1 - B(4; 25, .10) = .098$
 - c.** $P(X \geq 5) = 1 - B(4; 25, .20) = .579$
 - d.** All would decrease, which is bad if the % defective is large and good if the % is small.
- 97.**
- a.** $N = 500, p = .005$, so $np = 2.5$ and $b(x; 500, .005) \doteq p(x; 2.5)$, a Poisson p.m.f.
 - b.** $P(X = 5) = p(5; 2.5) - p(4; 2.5) = .9580 - .8912 = .0668$
 - c.** $P(X \geq 5) = 1 - p(4; 2.5) = 1 - .8912 = .1088$

Chapter 3: Discrete Random Variables and Probability Distributions

- 98.** $X \sim B(x; 25, p)$.
- $B(18; 25, .5) - B(6; 25, .5) = .986$
 - $B(18; 25, .8) - B(6; 25, .8) = .220$
 - With $p = .5$, $P(\text{rejecting the claim}) = P(X \leq 7) + P(X \geq 18) = .022 + [1 - .978] = .022 + .022 = .044$
 - The claim will not be rejected when $8 \leq X \leq 17$.
 With $p = .6$, $P(8 \leq X \leq 17) = B(17; 25, .6) - B(7; 25, .6) = .846 - .001 = .845$.
 With $p = .8$, $P(8 \leq X \leq 17) = B(17; 25, .8) - B(7; 25, .8) = .109 - .000 = .109$.
 - We want $P(\text{rejecting the claim}) \leq .01$. Using the decision rule “reject if $X \leq 6$ or $X \geq 19$ ” gives the probability .014, which is too large. We should use “reject if $X \leq 5$ or $X \geq 20$ ” which yields $P(\text{rejecting the claim}) = .002 + .002 = .004$.
- 99.** Let Y denote the number of tests carried out. For $n = 3$, possible Y values are 1 and 4. $P(Y = 1) = P(\text{no one has the disease}) = (.9)^3 = .729$ and $P(Y = 4) = .271$, so $E(Y) = (1)(.729) + (4)(.271) = 1.813$, as contrasted with the 3 tests necessary without group testing.
- 100.** Regard any particular symbol being received as constituting a trial. Then $p = P(S) = P(\text{symbol is sent correctly or is sent incorrectly and subsequently corrected}) = 1 - p_1 + p_1 p_2$. The block of n symbols gives a binomial experiment with n trials and $p = 1 - p_1 + p_1 p_2$.
- 101.** $p(2) = P(X = 2) = P(S \text{ on \#1 and } S \text{ on \#2}) = p^2$
 $p(3) = P(S \text{ on \#3 and } S \text{ on \#2 and } F \text{ on \#1}) = (1 - p)p^2$
 $p(4) = P(S \text{ on \#4 and } S \text{ on \#3 and } F \text{ on \#2}) = (1 - p)p^2$
 $p(5) = P(S \text{ on \#5 and } S \text{ on \#4 and } F \text{ on \#3 and no 2 consecutive S's on trials prior to \#3}) = [1 - p(2)](1 - p)p^2$
 $p(6) = P(S \text{ on \#6 and } S \text{ on \#5 and } F \text{ on \#4 and no 2 consecutive S's on trials prior to \#4}) = [1 - p(2) - p(3)](1 - p)p^2$
 In general, for $x = 5, 6, 7, \dots$: $p(x) = [1 - p(2) - \dots - p(x - 3)](1 - p)p^2$
 For $p = .9$,
- | x | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|------|-----|------|------|-------|-------|-------|-------|
| p(x) | .81 | .081 | .081 | .0154 | .0088 | .0023 | .0010 |
- So $P(X \leq 8) = p(2) + \dots + p(8) = .9995$
- 102.**
- With $X \sim \text{Bin}(25, .1)$, $P(2 \leq X \leq 6) = B(6; 25, .1) - B(1; 25, .1) = .991 - .271 = .720$
 - $E(X) = np = 25(.1) = 2.5$, $\sigma_X = \sqrt{npq} = \sqrt{25(.1)(.9)} = \sqrt{2.25} = 1.50$
 - $P(X \geq 7 \text{ when } p = .1) = 1 - B(6; 25, .1) = 1 - .991 = .009$
 - $P(X \leq 6 \text{ when } p = .2) = B(6; 25, .2) = .780$, which is quite large

Chapter 3: Discrete Random Variables and Probability Distributions

103.

- a. Let event C = seed carries single spikelets, and event P = seed produces ears with single spikelets. Then $P(P \cap C) = P(P | C) \cdot P(C) = .29(.40) = .116$. Let X = the number of seeds out of the 10 selected that meet the condition $P \cap C$. Then $X \sim \text{Bin}(10, .116)$.

$$P(X = 5) = \binom{10}{5} (.116)^5 (.884)^5 = .002857$$

- b. For 1 seed, the event of interest is P = seed produces ears with single spikelets.

$$P(P) = P(P \cap C) + P(P \cap C') = .116 \text{ (from a)} + P(P | C') \cdot P(C') \\ = .116 + (.26)(.40) = .272.$$

Let Y = the number out of the 10 seeds that meet condition P .

Then $Y \sim \text{Bin}(10, .272)$, and $P(Y = 5) = .0767$.

$$P(Y \leq 5) = b(0;10,.272) + \dots + b(5;10,.272) = .041813 + \dots + .076719 = .97024$$

104. With S = favored acquittal, the population size is $N = 12$, the number of population S 's is $M = 4$, the sample size is $n = 4$, and the p.m.f. of the number of interviewed jurors who favor

acquittal is the hypergeometric p.m.f. $h(x;4,4,12)$. $E(X) = 4 \cdot \left(\frac{4}{12}\right) = 1.33$

105.

- a. $P(X = 0) = F(0;2) = 0.135$

- b. Let S = an operator who receives no requests. Then $p = .135$ and we wish $P(4 \text{ S's in } 5 \text{ trials}) = b(4;5,.135) = \binom{5}{4} (.135)^4 (.884)^1 = .00144$

- c. $P(\text{all receive } x) = P(\text{first receives } x) \cdot \dots \cdot P(\text{fifth receives } x) = \left[\frac{e^{-2} 2^x}{x!} \right]^5$, and $P(\text{all receive the same number})$ is the sum from $x = 0$ to ∞ .

106. $P(\text{at least one}) = 1 - P(\text{none}) = 1 - e^{-\lambda \pi R^2} \cdot \frac{(\lambda \pi R^2)^0}{0!} = 1 - e^{-\lambda \pi R^2} = .99 \Rightarrow e^{-\lambda \pi R^2} = .01$

$$\Rightarrow R^2 = \frac{-\ln(.01)}{\lambda \pi} = .7329 \Rightarrow R = .8561$$

107. The number sold is $\min(X, 5)$, so $E[\min(x, 5)] = \sum_{x=0}^{\infty} \min(x, 5) p(x;4)$

$$= (0)p(0;4) + (1)p(1;4) + (2)p(2;4) + (3)p(3;4) + (4)p(4;4) + 5 \sum_{x=5}^{\infty} p(x;4)$$

$$= 1.735 + 5[1 - F(4;4)] = 3.59$$

Chapter 3: Discrete Random Variables and Probability Distributions

108.

- a. $P(X = x) = P(\text{A wins in } x \text{ games}) + P(\text{B wins in } x \text{ games})$
 $= P(9 \text{ S's in } 1^{\text{st}} x-1 \cap \text{S on the } x^{\text{th}}) + P(9 \text{ F's in } 1^{\text{st}} x-1 \cap \text{F on the } x^{\text{th}})$
 $= \binom{x-1}{9} p^9 (1-p)^{x-10} p + \binom{x-1}{9} (1-p)^9 p^{x-10} (1-p)$
 $= \binom{x-1}{9} [p^{10} (1-p)^{x-10} + (1-p)^{10} p^{x-10}]$
- b. Possible values of X are now 10, 11, 12, ... (all positive integers ≥ 10). Now
 $P(X = x) = \binom{x-1}{9} [p^{10} (1-p)^{x-10} + q^{10} (1-q)^{x-10}]$ for $x = 10, \dots, 19$,
 So $P(X \geq 20) = 1 - P(X < 20)$ and $P(X < 20) = \sum_{x=10}^{19} P(X = x)$

109.

- a. No; probability of success is not the same for all tests
- b. There are four ways exactly three could have positive results. Let D represent those with the disease and D' represent those without the disease.

Combination		Probability
D	D'	
0	3	$\left[\binom{5}{0} (.2)^0 (.8)^5 \right] \cdot \left[\binom{5}{3} (.9)^3 (.1)^2 \right]$ $= (.32768)(.0729) = .02389$
1	2	$\left[\binom{5}{1} (.2)^1 (.8)^4 \right] \cdot \left[\binom{5}{2} (.9)^2 (.1)^3 \right]$ $= (.4096)(.0081) = .00332$
2	1	$\left[\binom{5}{2} (.2)^2 (.8)^3 \right] \cdot \left[\binom{5}{1} (.9)^1 (.1)^4 \right]$ $= (.2048)(.00045) = .00009216$
3	0	$\left[\binom{5}{3} (.2)^3 (.8)^2 \right] \cdot \left[\binom{5}{0} (.9)^0 (.1)^5 \right]$ $= (.0512)(.00001) = .000000512$

Adding up the probabilities associated with the four combinations yields 0.0273.

Chapter 3: Discrete Random Variables and Probability Distributions

110. $k(r, x) = \frac{(x+r-1)(x+r-2)\dots(x+r-x)}{x!}$

With $r = 2.5$ and $p = .3$, $p(4) = \frac{(5.5)(4.5)(3.5)(2.5)}{4!} (.3)^{2.5} (.7)^4 = .1068$

Using $k(r, 0) = 1$, $P(X \geq 1) = 1 - p(0) = 1 - (.3)^{2.5} = .9507$

111.

- a.** $p(x; \lambda, \mu) = \frac{1}{2} p(x; \lambda) + \frac{1}{2} p(x; \mu)$ where both $p(x; \lambda)$ and $p(x; \mu)$ are Poisson p.m.f.'s and thus ≥ 0 , so $p(x; \lambda, \mu) \geq 0$. Further,

$$\sum_{x=0}^{\infty} p(x; \lambda, \mu) = \frac{1}{2} \sum_{x=0}^{\infty} p(x; \lambda) + \frac{1}{2} \sum_{x=0}^{\infty} p(x; \mu) = \frac{1}{2} + \frac{1}{2} = 1$$

- b.** $.6 p(x; \lambda) + .4 p(x; \mu)$

c. $E(X) = \sum_{x=0}^{\infty} x \left[\frac{1}{2} p(x; \lambda) + \frac{1}{2} p(x; \mu) \right] = \frac{1}{2} \sum_{x=0}^{\infty} x p(x; \lambda) + \frac{1}{2} \sum_{x=0}^{\infty} x p(x; \mu)$
 $= \frac{1}{2} \lambda + \frac{1}{2} \mu = \frac{\lambda + \mu}{2}$

d. $E(X^2) = \frac{1}{2} \sum_{x=0}^{\infty} x^2 p(x; \lambda) + \frac{1}{2} \sum_{x=0}^{\infty} x^2 p(x; \mu) = \frac{1}{2} (\lambda^2 + \lambda) + \frac{1}{2} (\mu^2 + \mu)$ (since for a Poisson r.v., $E(X^2) = V(X) + [E(X)]^2 = \lambda + \lambda^2$),
 so $V(X) = \frac{1}{2} [\lambda^2 + \lambda + \mu^2 + \mu] - \left[\frac{\lambda + \mu}{2} \right]^2 = \left(\frac{\lambda - \mu}{2} \right)^2 + \frac{\lambda + \mu}{2}$

112.

- a.** $\frac{b(x+1; n, p)}{b(x; n, p)} = \frac{(n-x)}{(x+1)} \cdot \frac{p}{(1-p)} > 1$ if $np - (1-p) > x$, from which the stated conclusion follows.

- b.** $\frac{p(x+1; \lambda)}{p(x; \lambda)} = \frac{\lambda}{(x+1)} > 1$ if $x < \lambda - 1$, from which the stated conclusion follows. If λ is an integer, then $\lambda - 1$ is a mode, but $p(\lambda, \lambda) = p(1 - \lambda, \lambda)$ so λ is also a mode [$p(x; \lambda)$] achieves its maximum for both $x = \lambda - 1$ and $x = \lambda$.

$$\begin{aligned}
 113. \quad P(X=j) &= \sum_{i=1}^{10} P(\text{arm on track } i \cap X=j) = \sum_{i=1}^{10} P(X=j \mid \text{arm on } i) \cdot p_i \\
 &= \sum_{i=1}^{10} P(\text{next seek at } I+j+1 \text{ or } I-j-1) \cdot p_i = \sum_{i=1}^{10} (p_{i+j+1} + p_{i-j-1}) p_i \\
 &\text{where } p_k = 0 \text{ if } k < 0 \text{ or } k > 10.
 \end{aligned}$$

$$\begin{aligned}
 114. \quad E(X) &= \sum_{x=0}^n x \cdot \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}} = \sum_{x=1}^n \frac{M!}{(x-1)!(M-x)!} \cdot \frac{\binom{N-M}{n-x}}{\binom{N}{n}} \\
 &= n \cdot \frac{M}{N} \sum_{x=1}^n \binom{M-1}{x-1} \frac{\binom{N-M}{n-x}}{\binom{N-1}{n-1}} = n \cdot \frac{M}{N} \sum_{y=0}^{n-1} \binom{M-1}{y} \frac{\binom{N-1-(M-1)}{n-1-y}}{\binom{N-1}{n-1}} \\
 &= n \cdot \frac{M}{N} \sum_{y=0}^{n-1} h(y; n-1, M-1, N-1) = n \cdot \frac{M}{N}
 \end{aligned}$$

$$\begin{aligned}
 115. \quad \text{Let } A = \{x: |x - \mu| \geq k\sigma\}. \text{ Then } \sigma^2 &= \sum_A (x - \mu)^2 p(x) \geq (k\sigma)^2 \sum_A p(x). \text{ But} \\
 \sum_A p(x) &= P(X \text{ is in } A) = P(|X - \mu| \geq k\sigma), \text{ so } \sigma^2 \geq k^2 \sigma^2 \cdot P(|X - \mu| \geq k\sigma), \text{ as desired.}
 \end{aligned}$$

116.

$$\text{a. For } [0,4], \lambda = \int_0^4 e^{2+.6t} dt = 123.44, \text{ whereas for } [2,6], \lambda = \int_2^6 e^{2+.6t} dt = 409.82$$

$$\text{b. } \lambda = \int_0^{0.9907} e^{2+.6t} dt = 9.9996 \approx 10, \text{ so the desired probability is } F(15, 10) = .951.$$

CHAPTER 4

Section 4.1

1.

$$\text{a. } P(x \leq 1) = \int_{-\infty}^1 f(x)dx = \int_0^1 \frac{1}{2}x dx = \frac{1}{4}x^2 \Big|_0^1 = .25$$

$$\text{b. } P(.5 \leq X \leq 1.5) = \int_{.5}^{1.5} \frac{1}{2}x dx = \frac{1}{4}x^2 \Big|_{.5}^{1.5} = .5$$

$$\text{c. } P(x > 1.5) = \int_{1.5}^{\infty} f(x)dx = \int_{1.5}^2 \frac{1}{2}x dx = \frac{1}{4}x^2 \Big|_{1.5}^2 = \frac{7}{16} \approx .438$$

2. $F(x) = \frac{1}{10}$ for $-5 \leq x \leq 5$, and $= 0$ otherwise

$$\text{a. } P(X < 0) = \int_{-5}^0 \frac{1}{10} dx = .5$$

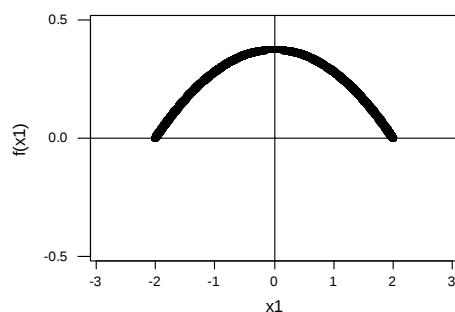
$$\text{b. } P(-2.5 < X < 2.5) = \int_{-2.5}^{2.5} \frac{1}{10} dx = .5$$

$$\text{c. } P(-2 \leq X \leq 3) = \int_{-2}^3 \frac{1}{10} dx = .5$$

$$\text{d. } P(k < X < k + 4) = \int_k^{k+4} \frac{1}{10} dx = \frac{x}{10} \Big|_k^{k+4} = \frac{1}{10}[(k + 4) - k] = .4$$

3.

a. Graph of $f(x) = .09375(4 - x^2)$



Chapter 4: Continuous Random Variables and Probability Distributions

$$\text{b. } P(X > 0) = \int_0^2 .09375(4 - x^2) dx = .09375 \left(4x - \frac{x^3}{3} \right) \Big|_0^2 = .5$$

$$\text{c. } P(-1 < X < 1) = \int_{-1}^1 .09375(4 - x^2) dx = .6875$$

$$\begin{aligned} \text{d. } P(X < -.5 \text{ OR } X > .5) &= 1 - P(-.5 \leq X \leq .5) = 1 - \int_{-.5}^{.5} .09375(4 - x^2) dx \\ &= 1 - .3672 = .6328 \end{aligned}$$

4.

$$\text{a. } \int_{-\infty}^{\infty} f(x; \theta) dx = \int_0^{\infty} \frac{x}{\theta^2} e^{-x^2/2\theta^2} dx = -e^{-x^2/2\theta^2} \Big|_0^{\infty} = 0 - (-1) = 1$$

$$\begin{aligned} \text{b. } P(X \leq 200) &= \int_{-\infty}^{200} f(x; \theta) dx = \int_0^{200} \frac{x}{\theta^2} e^{-x^2/2\theta^2} dx \\ &= -e^{-x^2/2\theta^2} \Big|_0^{200} \approx -.1353 + 1 = .8647 \end{aligned}$$

$P(X < 200) = P(X \leq 200) \approx .8647$, since x is continuous.

$P(X \geq 200) = 1 - P(X \leq 200) \approx .1353$

$$\text{c. } P(100 \leq X \leq 200) = \int_{100}^{200} f(x; \theta) dx = -e^{-x^2/20,000} \Big|_{100}^{200} \approx .4712$$

d. For $x > 0$, $P(X \leq x) =$

$$\int_{-\infty}^x f(y; \theta) dy = \int_0^x \frac{y}{\theta^2} e^{-y^2/2\theta^2} dy = -e^{-y^2/2\theta^2} \Big|_0^x = 1 - e^{-x^2/2\theta^2}$$

5.

$$\text{a. } 1 = \int_{-\infty}^{\infty} f(x) dx = \int_0^2 kx^2 dx = k \left(\frac{x^3}{3} \right) \Big|_0^2 = k \left(\frac{8}{3} \right) \Rightarrow k = \frac{3}{8}$$

$$\text{b. } P(0 \leq X \leq 1) = \int_0^1 \frac{3}{8} x^2 dx = \frac{1}{8} x^3 \Big|_0^1 = \frac{1}{8} = .125$$

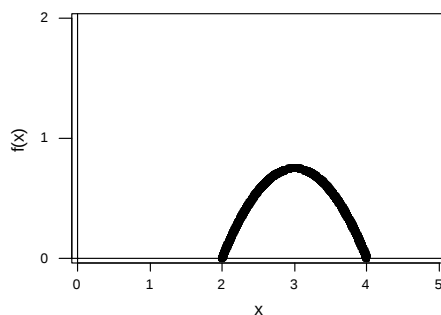
$$\text{c. } P(1 \leq X \leq 1.5) = \int_1^{1.5} \frac{3}{8} x^2 dx = \frac{1}{8} x^3 \Big|_1^{1.5} = \frac{1}{8} \left(\frac{3}{2} \right)^3 - \frac{1}{8} (1)^3 = \frac{19}{64} \approx .2969$$

$$\text{d. } P(X \geq 1.5) = 1 - \int_0^{1.5} \frac{3}{8} x^2 dx = \frac{1}{8} x^3 \Big|_0^{1.5} = 1 - \left[\frac{1}{8} \left(\frac{3}{2} \right)^3 - 0 \right] = 1 - \frac{27}{64} = \frac{37}{64} \approx .5781$$

Chapter 4: Continuous Random Variables and Probability Distributions

6.

a.



$$\text{b. } 1 = \int_2^4 k[1 - (x-3)^2] dx = \int_{-1}^1 k[1 - u^2] du = \frac{4}{3} \Rightarrow k = \frac{3}{4}$$

$$\text{c. } P(X > 3) = \int_3^4 \frac{3}{4}[1 - (x-3)^2] dx = .5 \text{ by symmetry of the p.d.f}$$

$$\text{d. } P\left(\frac{11}{4} \leq X \leq \frac{13}{4}\right) = \int_{11/4}^{13/4} \frac{3}{4}[1 - (x-3)^2] dx = \frac{3}{4} \int_{-1/4}^{1/4} [1 - (u)^2] du = \frac{47}{128} \approx .367$$

$$\begin{aligned} \text{e. } P(|X-3| > .5) &= 1 - P(|X-3| \leq .5) = 1 - P(2.5 \leq X \leq 3.5) \\ &= 1 - \int_{-5}^5 \frac{3}{4}[1 - (u)^2] du = \frac{5}{16} \approx .313 \end{aligned}$$

7.

$$\text{a. } f(x) = \frac{1}{10} \text{ for } 25 \leq x \leq 35 \text{ and } = 0 \text{ otherwise}$$

$$\text{b. } P(X > 33) = \int_{33}^{35} \frac{1}{10} dx = .2$$

$$\text{c. } E(X) = \int_{25}^{35} x \cdot \frac{1}{10} dx = \left. \frac{x^2}{20} \right|_{25}^{35} = 30$$

30 ± 2 is from 28 to 32 minutes:

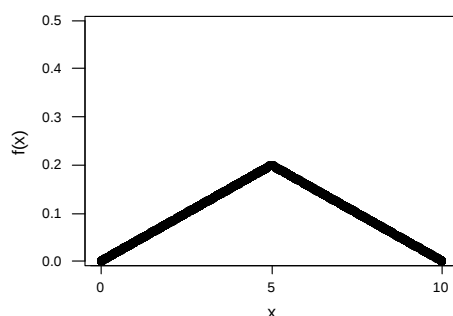
$$P(28 < X < 32) = \int_{28}^{32} \frac{1}{10} dx = \left. \frac{1}{10} x \right|_{28}^{32} = .4$$

$$\text{d. } P(a \leq x \leq a+2) = \int_a^{a+2} \frac{1}{10} dx = .2, \text{ since the interval has length 2.}$$

Chapter 4: Continuous Random Variables and Probability Distributions

8.

a.



$$\begin{aligned} \text{b. } \int_{-\infty}^{\infty} f(y)dy &= \int_0^5 \frac{1}{25}ydy + \int_5^{10} \left(\frac{2}{5} - \frac{1}{25}y\right)dy = \left[\frac{y^2}{50}\right]_0^5 + \left[\frac{2}{5}y - \frac{1}{50}y^2\right]_5^{10} \\ &= \frac{1}{2} + \left[(4 - 2) - \left(2 - \frac{1}{2}\right)\right] = \frac{1}{2} + \frac{1}{2} = 1 \end{aligned}$$

$$\text{c. } P(Y \leq 3) = \int_0^3 \frac{1}{25}ydy = \left[\frac{y^2}{50}\right]_0^3 = \frac{9}{50} \approx .18$$

$$\text{d. } P(Y \leq 8) = \int_0^5 \frac{1}{25}ydy + \int_5^8 \left(\frac{2}{5} - \frac{1}{25}y\right)dy = \frac{23}{25} \approx .92$$

$$\text{e. } P(3 \leq Y \leq 8) = P(Y \leq 8) - P(Y < 3) = \frac{46}{50} - \frac{9}{50} = \frac{37}{50} = .74$$

$$\text{f. } P(Y < 2 \text{ or } Y > 6) = \int_0^2 \frac{1}{25}ydy + \int_6^{10} \left(\frac{2}{5} - \frac{1}{25}y\right)dy = \frac{2}{5} = .4$$

9.

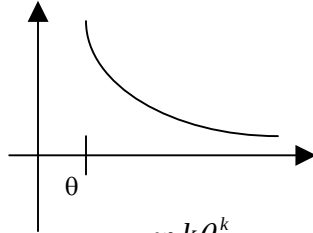
$$\begin{aligned} \text{a. } P(X \leq 6) &= \int_{.5}^6 .15e^{-.15(x-.5)}dx = .15 \int_0^{5.5} e^{-.15u}du \quad (\text{after } u = x - .5) \\ &= \left[e^{-.15u}\right]_0^{5.5} = 1 - e^{-.825} \approx .562 \end{aligned}$$

$$\text{b. } 1 - .562 = .438; .438$$

$$\text{c. } P(5 \leq Y \leq 6) = P(Y \leq 6) - P(Y \leq 5) \approx .562 - .491 = .071$$

10.

a.



$$\text{b. } \int_{-\infty}^{\infty} f(x; k, \theta) dx = \int_{\theta}^{\infty} \frac{k\theta^k}{x^{k+1}} dx = \theta^k \cdot \left(-\frac{1}{x^k} \right) \Big|_{\theta}^{\infty} = \frac{\theta^k}{\theta^k} = 1$$

$$\text{c. } P(X \leq b) = \int_{\theta}^b \frac{k\theta^k}{x^{k+1}} dx = \theta^k \cdot \left(-\frac{1}{x^k} \right) \Big|_{\theta}^b = 1 - \left(\frac{\theta}{b} \right)^k$$

$$\text{d. } P(a \leq X \leq b) = \int_a^b \frac{k\theta^k}{x^{k+1}} dx = \theta^k \cdot \left(-\frac{1}{x^k} \right) \Big|_a^b = \left(\frac{\theta}{a} \right)^k - \left(\frac{\theta}{b} \right)^k$$

Section 4.2

11.

$$\text{a. } P(X \leq 1) = F(1) = \frac{1}{4} = .25$$

$$\text{b. } P(.5 \leq X \leq 1) = F(1) - F(.5) = \frac{3}{16} = .1875$$

$$\text{c. } P(X > .5) = 1 - P(X \leq .5) = 1 - F(.5) = \frac{15}{16} = .9375$$

$$\text{d. } .5 = F(\tilde{\mu}) = \frac{\tilde{\mu}^2}{4} \Rightarrow \tilde{\mu}^2 = 2 \Rightarrow \tilde{\mu} = \sqrt{2} \approx 1.414$$

$$\text{e. } f(x) = f'(x) = \frac{x}{2} \text{ for } 0 \leq x < 2, \text{ and } = 0 \text{ otherwise}$$

$$\text{f. } E(X) = \int_{-\infty}^{\infty} x \cdot f(x) dx = \int_0^2 x \cdot \frac{1}{2} x dx = \frac{1}{2} \int_0^2 x^2 dx = \frac{x^3}{6} \Big|_0^2 = \frac{8}{6} \approx 1.333$$

$$\text{g. } E(X^2) = \int_{-\infty}^{\infty} x^2 f(x) dx = \int_0^2 x^2 \cdot \frac{1}{2} x dx = \frac{1}{2} \int_0^2 x^3 dx = \frac{x^4}{8} \Big|_0^2 = 2,$$

$$\text{So } \text{Var}(X) = E(X^2) - [E(X)]^2 = 2 - \left(\frac{8}{6} \right)^2 = \frac{8}{36} \approx .222, \sigma_x \approx .471$$

$$\text{h. } \text{From g, } E(X^2) = 2$$

12.

a. $P(X < 0) = F(0) = .5$

b. $P(-1 \leq X \leq 1) = F(1) - F(-1) = \frac{11}{16} = .6875$

c. $P(X > .5) = 1 - P(X \leq .5) = 1 - F(.5) = 1 - .6836 = .3164$

d. $F(x) = F'(x) = \frac{d}{dx} \left(\frac{1}{2} + \frac{3}{32} \left(4x - \frac{x^3}{3} \right) \right) = 0 + \frac{3}{32} \left(4 - \frac{3x^2}{3} \right) = .09375(4 - x^2)$

e. $F(\tilde{\mu}) = .5$ by definition. $F(0) = .5$ from a above, which is as desired.

13.

a. $1 = \int_1^{\infty} \frac{k}{x^4} dx \Rightarrow 1 = \frac{-k}{3} X^{-3} \Big|_1^{\infty} \Rightarrow 1 = 0 - \left(-\frac{k}{3}\right)(1) \Rightarrow 1 = \frac{k}{3} \Rightarrow k = 3$

b. cdf: $F(x) = \int_{-\infty}^x f(y) dy = \int_1^x 3y^{-4} dy = -\frac{3}{3} y^{-3} \Big|_1^x = -x^{-3} + 1 = 1 - \frac{1}{x^3}$. So

$$F(x) = \begin{cases} 0, & x \leq 1 \\ 1 - x^{-3}, & x > 1 \end{cases}$$

c. $P(x > 2) = 1 - F(2) = 1 - \left(1 - \frac{1}{8}\right) = \frac{1}{8}$ or .125;
 $P(2 < x < 3) = F(3) - F(2) = \left(1 - \frac{1}{27}\right) - \left(1 - \frac{1}{8}\right) = .963 - .875 = .088$

d. $E(x) = \int_1^{\infty} x \left(\frac{3}{x^4} \right) dx = \int_1^{\infty} \left(\frac{3}{x^3} \right) dx = -\frac{3}{2} X^{-2} \Big|_1^{\infty} = 0 + \frac{3}{2} = \frac{3}{2}$
 $E(x^2) = \int_1^{\infty} x^2 \left(\frac{3}{x^4} \right) dx = \int_1^{\infty} \left(\frac{3}{x^2} \right) dx = -3X^{-1} \Big|_1^{\infty} = 0 + 3 = 3$
 $V(x) = E(x^2) - [E(x)]^2 = 3 - \left(\frac{3}{2}\right)^2 = 3 - \frac{9}{4} = \frac{3}{4}$ or .75
 $\sigma = \sqrt{V(x)} = \sqrt{\frac{3}{4}} = .866$

e. $P(1.5 - .866 < x < 1.5 + .866) = P(x < 2.366) = F(2.366)$
 $= 1 - (2.366^{-3}) = .9245$

14.

- a. If X is uniformly distributed on the interval from A to B , then

$$E(X) = \int_A^B x \cdot \frac{1}{B-A} dx = \frac{A+B}{2}, E(X^2) = \frac{A^2 + AB + B^2}{3}$$

$$V(X) = E(X^2) - [E(X)]^2 = \frac{(B-A)^2}{12}$$

With $A = 7.5$ and $B = 20$, $E(X) = 13.75$, $V(X) = 13.02$

b.
$$F(X) = \begin{cases} 0 & x < 7.5 \\ \frac{x-7.5}{12.5} & 7.5 \leq x < 20 \\ 1 & x \geq 20 \end{cases}$$

- c. $P(X \leq 10) = F(10) = .200$; $P(10 \leq X \leq 15) = F(15) - F(10) = .4$

- d. $\sigma = 3.61$, so $\mu \pm \sigma = (10.14, 17.36)$

Thus, $P(\mu - \sigma \leq X \leq \mu + \sigma) = F(17.36) - F(10.14) = .5776$

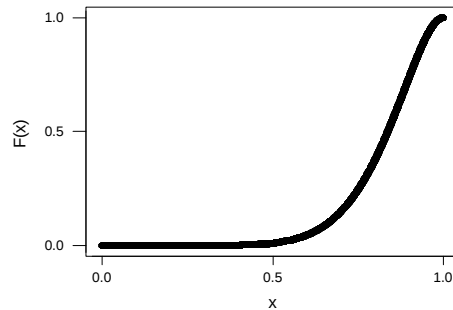
Similarly, $P(\mu - \sigma \leq X \leq \mu + \sigma) = P(6.53 \leq X \leq 20.97) = 1$

15.

- a. $F(X) = 0$ for $x \leq 0$, $= 1$ for $x \geq 1$, and for $0 < X < 1$,

$$F(X) = \int_{-\infty}^x f(y) dy = \int_0^x 90y^8(1-y) dy = 90 \int_0^x (y^8 - y^9) dy$$

$$90 \left(\frac{1}{9} y^9 - \frac{1}{10} y^{10} \right) \Big|_0^x = 10x^9 - 9x^{10}$$



- b. $F(.5) = 10(.5)^9 - 9(.5)^{10} \approx .0107$

- c. $P(.25 \leq X \leq .5) = F(.5) - F(.25) \approx .0107 - [10(.25)^9 - 9(.25)^{10}]$
 $\approx .0107 - .0000 \approx .0107$

- d. The 75th percentile is the value of x for which $F(x) = .75$
 $\Rightarrow .75 = 10(x)^9 - 9(x)^{10} \Rightarrow x \approx .9036$

Chapter 4: Continuous Random Variables and Probability Distributions

$$\begin{aligned} \text{e. } E(X) &= \int_{-\infty}^{\infty} x \cdot f(x) dx = \int_0^1 x \cdot 90x^8(1-x) dx = 90 \int_0^1 x^9(1-x) dx \\ &= 9x^{10} - \frac{90}{11}x^{11} \Big|_0^1 = \frac{9}{11} \approx .8182 \\ E(X^2) &= \int_{-\infty}^{\infty} x^2 \cdot f(x) dx = \int_0^1 x^2 \cdot 90x^8(1-x) dx = 90 \int_0^1 x^{10}(1-x) dx \\ &= \frac{90}{11}x^{11} - \frac{90}{12}x^{12} \Big|_0^1 \approx .6818 \end{aligned}$$

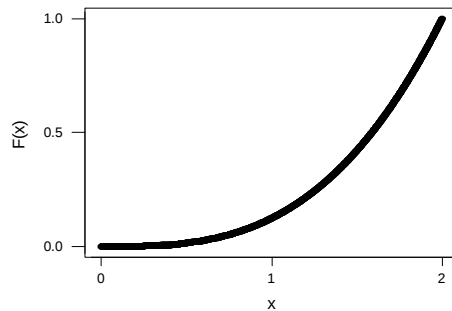
$$V(X) \approx .6818 - (.8182)^2 = .0124, \quad \sigma_x = .11134.$$

$$\begin{aligned} \text{f. } \mu \pm \sigma &= (.7068, .9295). \text{ Thus, } P(\mu - \sigma \leq X \leq \mu + \sigma) = F(.9295) - F(.7068) \\ &= .8465 - .1602 = .6863 \end{aligned}$$

16.

a. $F(x) = 0$ for $x < 0$ and $F(x) = 1$ for $x > 2$. For $0 \leq x \leq 2$,

$$F(x) = \int_0^x \frac{3}{8}y^2 dy = \frac{1}{8}y^3 \Big|_0^x = \frac{1}{8}x^3$$



$$\text{b. } P(x \leq .5) = F(.5) = \frac{1}{8} \left(\frac{1}{2} \right)^3 = \frac{1}{64}$$

$$\text{c. } P(.25 \leq X \leq .5) = F(.5) - F(.25) = \frac{1}{64} - \frac{1}{8} \left(\frac{1}{4} \right)^3 = \frac{7}{512} \approx .0137$$

$$\text{d. } .75 = F(x) = \frac{1}{8}x^3 \Rightarrow x^3 = 6 \Rightarrow x \approx 1.8171$$

$$\text{e. } E(X) = \int_{-\infty}^{\infty} x \cdot f(x) dx = \int_0^2 x \cdot \left(\frac{3}{8}x^2 \right) dx = \frac{3}{8} \int_0^1 x^3 dx = \frac{3}{8} \left(\frac{1}{4}x^4 \right) \Big|_0^2 = \frac{3}{2} = 1.5$$

$$E(X^2) = \int_0^2 x \cdot \left(\frac{3}{8}x^2 \right) dx = \frac{3}{8} \int_0^1 x^4 dx = \frac{3}{8} \left(\frac{1}{5}x^5 \right) \Big|_0^2 = \frac{12}{5} = 2.4$$

$$V(X) = \frac{12}{5} - \left(\frac{3}{2} \right)^2 = \frac{3}{20} = .15 \quad \sigma_x = .3873$$

$$\text{f. } \mu \pm \sigma = (1.1127, 1.8873). \text{ Thus, } P(\mu - \sigma \leq X \leq \mu + \sigma) = F(1.8873) - F(1.1127) = .8403 - .1722 = .6681$$

17.

a. For $2 \leq X \leq 4$, $F(X) = \int_{-\infty}^x f(y)dy = \int_2^x \frac{3}{4}[1 - (y-3)^2]dy$ (let $u = y-3$)

$$= \int_{-1}^{x-3} \frac{3}{4}[1 - u^2]du = \frac{3}{4} \left[u - \frac{u^3}{3} \right]_{-1}^{x-3} = \frac{3}{4} \left[x - \frac{7}{3} - \frac{(x-3)^3}{3} \right]. \text{ Thus}$$

$$F(x) = \begin{cases} 0 & x < 2 \\ \frac{1}{4}[3x - 7 - (x-3)^3] & 2 \leq x \leq 4 \\ 1 & x > 4 \end{cases}$$

b. By symmetry of $f(x)$, $\tilde{\mu} = 3$

c. $E(X) = \int_2^4 x \cdot \frac{3}{4}[1 - (x-3)^2]dx = \frac{3}{4} \int_{-1}^1 (y+3)(1-y^2)dy$

$$= \frac{3}{4} \left[3y + \frac{y^2}{2} - y^3 - \frac{y^4}{4} \right]_{-1}^1 = \frac{3}{4} \cdot 4 = 3$$

$$V(X) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx = \frac{3}{4} \int_2^4 (x-3)^2 \cdot [1 - (x-3)^2]dx$$

$$= \frac{3}{4} \int_{-1}^1 y^2(1-y^2)dy = \frac{3}{4} \cdot \frac{4}{15} = \frac{1}{5} = .2$$

18.

a. $F(X) = \frac{x-A}{B-A} = p \Rightarrow x = (100p)\text{th percentile} = A + (B-A)p$

b. $E(X) = \int_A^B x \cdot \frac{1}{B-A} dx = \frac{1}{B-A} \cdot \frac{x^2}{2} \Big|_A^B = \frac{1}{2} \cdot \frac{1}{B-A} \cdot (B^2 - A^2) = \frac{A+B}{2}$

$$E(X^2) = \frac{1}{3} \cdot \frac{1}{B-A} \cdot (B^3 - A^3) = \frac{A^2 + AB + B^2}{3}$$

$$V(X) = \left(\frac{A^2 + AB + B^2}{3} \right) - \left(\frac{(A+B)^2}{4} \right) = \frac{(B-A)^2}{12}, \quad \sigma_x = \frac{(B-A)}{\sqrt{12}}$$

c. $E(X^n) = \int_A^B x^n \cdot \frac{1}{B-A} dx = \frac{B^{n+1} - A^{n+1}}{(n+1)(B-A)}$

Chapter 4: Continuous Random Variables and Probability Distributions

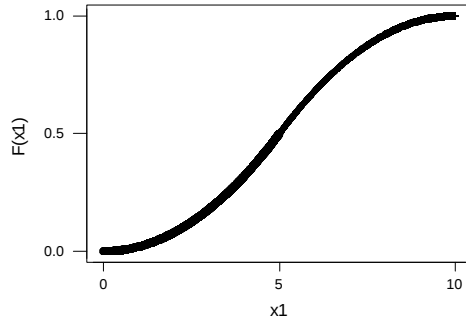
19.

- a. $P(X \leq 1) = F(1) = .25[1 + \ln(4)] \approx .597$
- b. $P(1 \leq X \leq 3) = F(3) - F(1) \approx .966 - .597 \approx .369$
- c. $f(x) = F'(x) = .25 \ln(4) - .25 \ln(x)$ for $0 < x < 4$

20.

- a. For $0 \leq y \leq 5$, $F(y) = \int_0^y \frac{1}{25} u du = \frac{y^2}{50}$
 For $5 \leq y \leq 10$, $F(y) = \int_0^y f(u) du = \int_0^5 f(u) du + \int_5^y f(u) du$

$$= \frac{1}{2} + \int_5^y \left(\frac{2}{5} - \frac{u}{25} \right) du = \frac{2}{5} y - \frac{y^2}{50} - 1$$



- b. For $0 < p \leq .5$, $p = F(y_p) = \frac{y_p^2}{50} \Rightarrow y_p = (50p)^{1/2}$
 For $.5 < p \leq 1$, $p = \frac{2}{5} y_p - \frac{y_p^2}{50} - 1 \Rightarrow y_p = 10 - 5\sqrt{2(1-p)}$
- c. $E(Y) = 5$ by straightforward integration (or by symmetry of $f(y)$), and similarly $V(Y) = \frac{50}{12} = 4.1667$. For the waiting time X for a single bus,
 $E(X) = 2.5$ and $V(X) = \frac{25}{12}$

21.
$$E(\text{area}) = E(\pi R^2) = \int_{-\infty}^{\infty} \pi r^2 f(r) dr = \int_9^{11} \pi r^2 \left(\frac{3}{4} \right) (1 - (10 - r)^2) dr$$

$$= \left(\frac{3}{4} \right) \pi \int_9^{11} r^2 (1 - (100 - 20r + r^2)) dr = \frac{3}{4} \pi \int_9^{11} -99r^2 + 20r^3 - r^4 dr = 100 \cdot 2\pi$$

22.

a. For $1 \leq x \leq 2$, $F(x) = \int_1^x 2\left(1 - \frac{1}{y^2}\right) dy = 2\left(y + \frac{1}{y}\right) \Big|_1^x = 2\left(x + \frac{1}{x}\right) - 4$, so

$$F(x) = \begin{cases} 0 & x < 1 \\ 2\left(x + \frac{1}{x}\right) - 4 & 1 \leq x \leq 2 \\ 1 & x > 2 \end{cases}$$

b. $2\left(x_p + \frac{1}{x_p}\right) - 4 = p \Rightarrow 2x_p^2 - (4-p)x_p + 2 = 0 \Rightarrow x_p = \frac{1}{4}[4 + p + \sqrt{p^2 + 8p}]$ To find $\tilde{\mu}$, set $p = .5 \Rightarrow \tilde{\mu} = 1.64$

c. $E(X) = \int_1^2 x \cdot 2\left(1 - \frac{1}{x^2}\right) dx = 2 \int_1^2 \left(x - \frac{1}{x}\right) dx = 2\left(\frac{x^2}{2} - \ln(x)\right) \Big|_1^2 = 1.614$

$$E(X^2) = 2 \int_1^2 (x^2 - 1) dx = 2\left(\frac{x^3}{3} - x\right) \Big|_1^2 = \frac{8}{3} \Rightarrow \text{Var}(X) = .0626$$

d. Amount left = $\max(1.5 - X, 0)$, so

$$E(\text{amount left}) = \int_1^2 \max(1.5 - x, 0) f(x) dx = 2 \int_1^{1.5} (1.5 - x) \left(1 - \frac{1}{x^2}\right) dx = .061$$

23. With X = temperature in $^{\circ}\text{C}$, temperature in $^{\circ}\text{F} = \frac{9}{5}X + 32$, so

$$E\left[\frac{9}{5}X + 32\right] = \frac{9}{5}(120) + 32 = 248, \quad \text{Var}\left[\frac{9}{5}X + 32\right] = \left(\frac{9}{5}\right)^2 \cdot (2)^2 = 12.96,$$

so $\sigma = 3.6$

24.

$$\text{a. } E(X) = \int_{\theta}^{\infty} x \cdot \frac{k\theta^k}{x^{k+1}} dx = k\theta^k \int_{\theta}^{\infty} \frac{1}{x^k} dx = \frac{k\theta^k x^{-k+1}}{-k+1} \Big|_{\theta}^{\infty} = \frac{k\theta}{k-1}$$

$$\text{b. } E(X) = \infty$$

$$\text{c. } E(X^2) = k\theta^k \int_{\theta}^{\infty} \frac{1}{x^{k-1}} dx = \frac{k\theta^2}{k-2}, \text{ so}$$

$$\text{Var}(X) = \left(\frac{k\theta^2}{k-2} \right) - \left(\frac{k\theta}{k-1} \right)^2 = \frac{k\theta^2}{(k-2)(k-1)^2}$$

$$\text{d. } \text{Var}(x) = \infty, \text{ since } E(X^2) = \infty.$$

$$\text{e. } E(X^n) = k\theta^k \int_{\theta}^{\infty} x^{n-(k+1)} dx, \text{ which will be finite if } n - (k+1) < -1, \text{ i.e. if } n < k.$$

25.

$$\text{a. } P(Y \leq 1.8\tilde{\mu} + 32) = P(1.8X + 32 \leq 1.8\tilde{\mu} + 32) = P(X \leq \tilde{\mu}) = .5$$

$$\text{b. } 90^{\text{th}} \text{ for } Y = 1.8\eta(.9) + 32 \text{ where } \eta(.9) \text{ is the } 90^{\text{th}} \text{ percentile for } X, \text{ since}$$

$$P(Y \leq 1.8\eta(.9) + 32) = P(1.8X + 32 \leq 1.8\eta(.9) + 32)$$

$$= P(X \leq \eta(.9)) = .9 \text{ as desired.}$$

$$\text{c. } \text{The } (100p)\text{th percentile for } Y \text{ is } 1.8\eta(p) + 32, \text{ verified by substituting } p \text{ for } .9 \text{ in the}$$

argument of **b**. When $Y = aX + b$, (i.e. a linear transformation of X), and the $(100p)$ th percentile of the X distribution is $\eta(p)$, then the corresponding $(100p)$ th percentile of the Y distribution is $a\eta(p) + b$. (same linear transformation applied to X 's percentile)

Section 4.3

26.

- a. $P(0 \leq Z \leq 2.17) = \Phi(2.17) - \Phi(0) = .4850$
- b. $\Phi(1) - \Phi(0) = .3413$
- c. $\Phi(0) - \Phi(-2.50) = .4938$
- d. $\Phi(2.50) - \Phi(-2.50) = .9876$
- e. $\Phi(1.37) = .9147$
- f. $P(-1.75 < Z) + [1 - P(Z < -1.75)] = 1 - \Phi(-1.75) = .9599$
- g. $\Phi(2) - \Phi(-1.50) = .9104$
- h. $\Phi(2.50) - \Phi(1.37) = .0791$
- i. $1 - \Phi(1.50) = .0668$
- j. $P(|Z| \leq 2.50) = P(-2.50 \leq Z \leq 2.50) = \Phi(2.50) - \Phi(-2.50) = .9876$

27.

- a. .9838 is found in the 2.1 row and the .04 column of the standard normal table so $c = 2.14$.
- b. $P(0 \leq Z \leq c) = .291 \Rightarrow \Phi(c) = .7910 \Rightarrow c = .81$
- c. $P(c \leq Z) = .121 \Rightarrow 1 - P(c \leq Z) = P(Z < c) = \Phi(c) = 1 - .121 = .8790 \Rightarrow c = 1.17$
- d. $P(-c \leq Z \leq c) = \Phi(c) - \Phi(-c) = \Phi(c) - (1 - \Phi(c)) = 2\Phi(c) - 1$
 $\Rightarrow \Phi(c) = .9920 \Rightarrow c = .97$
- e. $P(c \leq |Z|) = .016 \Rightarrow 1 - .016 = .9840 = 1 - P(c \leq |Z|) = P(|Z| < c)$
 $= P(-c < Z < c) = \Phi(c) - \Phi(-c) = 2\Phi(c) - 1$
 $\Rightarrow \Phi(c) = .9920 \Rightarrow c = 2.41$

Chapter 4: Continuous Random Variables and Probability Distributions

28.

- a. $\Phi(c) = .9100 \Rightarrow c \approx 1.34$ (.9099 is the entry in the 1.3 row, .04 column)
- b. 9th percentile = -91st percentile = -1.34
- c. $\Phi(c) = .7500 \Rightarrow c \approx .675$ since .7486 and .7517 are in the .67 and .68 entries, respectively.
- d. $25^{\text{th}} = -75^{\text{th}} = -.675$
- e. $\Phi(c) = .06 \Rightarrow c \approx -1.555$ (both .0594 and .0606 appear as the -1.56 and -1.55 entries, respectively).

29.

- a. Area under Z curve above $z_{.0055}$ is .0055, which implies that $\Phi(z_{.0055}) = 1 - .0055 = .9945$, so $z_{.0055} = 2.54$
- b. $\Phi(z_{.09}) = .9100 \Rightarrow z = 1.34$ (since .9099 appears as the 1.34 entry).
- c. $\Phi(z_{.633}) = \text{area below } z_{.633} = .3370 \Rightarrow z_{.633} \approx -.42$

30.

- a. $P(X \leq 100) = P\left(z \leq \frac{100 - 80}{10}\right) = P(Z \leq 2) = \Phi(2.00) = .9772$
- b. $P(X \leq 80) = P\left(z \leq \frac{80 - 80}{10}\right) = P(Z \leq 0) = \Phi(0.00) = .5$
- c. $P(65 \leq X \leq 100) = P\left(\frac{65 - 80}{10} \leq z \leq \frac{100 - 80}{10}\right) = P(-1.50 \leq Z \leq 2)$
 $= \Phi(2.00) - \Phi(-1.50) = .9772 - .0668 = .9104$
- d. $P(70 \leq X) = P(-1.00 \leq Z) = 1 - \Phi(-1.00) = .8413$
- e. $P(85 \leq X \leq 95) = P(.50 \leq Z \leq 1.50) = \Phi(1.50) - \Phi(.50) = .2417$
- f. $P(|X - 80| \leq 10) = P(-10 \leq X - 80 \leq 10) = P(70 \leq X \leq 90)$
 $P(-1.00 \leq Z \leq 1.00) = .6826$

Chapter 4: Continuous Random Variables and Probability Distributions

31.

- a. $P(X \leq 18) = P\left(Z \leq \frac{18 - 15}{1.25}\right) = P(Z \leq 2.4) = \Phi(2.4) = .9452$
- b. $P(10 \leq X \leq 12) = P(-4.00 \leq Z \leq -2.40) \approx P(Z \leq -2.40) = \Phi(-2.40) = .0082$
- c. $P(|X - 10| \leq 2(1.25)) = P(-2.50 \leq X - 15 \leq 2.50) = P(12.5 \leq X \leq 17.5)$
 $P(-2.00 \leq Z \leq 2.00) = .9544$

32.

- a. $P(X > .25) = P(Z > -.83) = 1 - .2033 = .7967$
- b. $P(X \leq .10) = \Phi(-3.33) = .0004$
- c. We want the value of the distribution, c , that is the 95th percentile (5% of the values are higher). The 95th percentile of the standard normal distribution = 1.645. So $c = .30 + (1.645)(.06) = .3987$. The largest 5% of all concentration values are above .3987 mg/cm³.

33.

- a. $P(X \geq 10) = P(Z \geq .43) = 1 - \Phi(.43) = 1 - .6664 = .3336$.
 $P(X > 10) = P(X \geq 10) = .3336$, since for any continuous distribution, $P(x = a) = 0$.
- b. $P(X > 20) = P(Z > 4) \approx 0$
- c. $P(5 \leq X \leq 10) = P(-1.36 \leq Z \leq .43) = \Phi(.43) - \Phi(-1.36) = .6664 - .0869 = .5795$
- d. $P(8.8 - c \leq X \leq 8.8 + c) = .98$, so $8.8 - c$ and $8.8 + c$ are at the 1st and the 99th percentile of the given distribution, respectively. The 1st percentile of the standard normal distribution has the value -2.33 , so
 $8.8 - c = \mu + (-2.33)\sigma = 8.8 - 2.33(2.8) \Rightarrow c = 2.33(2.8) = 6.524$.
- e. From a, $P(x > 10) = .3336$. Define event A as {diameter > 10 }, then $P(\text{at least one } A_i) = 1 - P(\text{no } A_i) = 1 - P(A')^4 = 1 - (1 - .3336)^4 = 1 - .1972 = .8028$

34.

Let X denote the diameter of a randomly selected cork made by the first machine, and let Y be defined analogously for the second machine.
 $P(2.9 \leq X \leq 3.1) = P(-1.00 \leq Z \leq 1.00) = .6826$
 $P(2.9 \leq Y \leq 3.1) = P(-7.00 \leq Z \leq 3.00) = .9987$
 So the second machine wins handily.

Chapter 4: Continuous Random Variables and Probability Distributions

35.

a. $\mu + \sigma \cdot (91^{\text{st}} \text{ percentile from std normal}) = 30 + 5(1.34) = 36.7$

b. $30 + 5(-1.555) = 22.225$

c. $\mu = 3.000 \mu\text{m}; \sigma = 0.140$. We desire the 90th percentile: $30 + 1.28(0.14) = 3.179$

36.

$\mu = 43; \sigma = 4.5$

a. $P(X < 40) = P\left(z \leq \frac{40 - 43}{4.5}\right) = P(Z < -0.667) = .2514$

$$P(X > 60) = P\left(z > \frac{60 - 43}{4.5}\right) = P(Z > 3.778) \approx 0$$

b. $43 + (-0.67)(4.5) = 39.985$

37.

$$P(\text{damage}) = P(X < 100) = P\left(z < \frac{100 - 200}{300}\right) = P(Z < -3.33) = .0004$$

$$\begin{aligned} P(\text{at least one among five is damaged}) &= 1 - P(\text{none damaged}) \\ &= 1 - (.9996)^5 = 1 - .998 = .002 \end{aligned}$$

38.

From Table A.3, $P(-1.96 \leq Z \leq 1.96) = .95$. Then $P(\mu - .1 \leq X \leq \mu + .1) =$

$$P\left(\frac{-.1}{\sigma} < z < \frac{.1}{\sigma}\right) \text{ implies that } \frac{.1}{\sigma} = 1.96, \text{ and thus that } \sigma = \frac{.1}{1.96} = .0510$$

39.

Since 1.28 is the 90th z percentile ($z_{.1} = 1.28$) and -1.645 is the 5th z percentile ($z_{.05} = 1.645$), the given information implies that $\mu + \sigma(1.28) = 10.256$ and $\mu + \sigma(-1.645) = 9.671$, from which $\sigma(-2.925) = -.585$, $\sigma = .2000$, and $\mu = 10$.

40.

a. $P(\mu - 1.5\sigma \leq X \leq \mu + 1.5\sigma) = P(-1.5 \leq Z \leq 1.5) = \Phi(1.50) - \Phi(-1.50) = .8664$

b. $P(X < \mu - 2.5\sigma \text{ or } X > \mu + 2.5\sigma) = 1 - P(\mu - 2.5\sigma \leq X \leq \mu + 2.5\sigma)$
 $= 1 - P(-2.5 \leq Z \leq 2.5) = 1 - .9876 = .0124$

c. $P(\mu - 2\sigma \leq X \leq \mu - \sigma \text{ or } \mu + \sigma \leq X \leq \mu + 2\sigma) = P(\text{within 2 sd's}) - P(\text{within 1 sd}) = P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) - P(\mu - \sigma \leq X \leq \mu + \sigma)$
 $= .9544 - .6826 = .2718$

Chapter 4: Continuous Random Variables and Probability Distributions

41. With $\mu = .500$ inches, the acceptable range for the diameter is between .496 and .504 inches, so unacceptable bearings will have diameters smaller than .496 or larger than .504. The new distribution has $\mu = .499$ and $\sigma = .002$. $P(x < .496 \text{ or } x > .504) =$

$$P\left(z < \frac{.496 - .499}{.002}\right) + P\left(z > \frac{.504 - .499}{.002}\right) = P(z < -1.5) + P(z > 2.5)$$

$\Phi(-1.5) + (1 - \Phi(2.5)) = .0068 + .0062 = .013$, or 1.3% of the bearings will be unacceptable.

42.

a. $P(67 \leq X \leq 75) = P(-1.00 \leq Z \leq 1.67) = .7938$

b. $P(70 - c \leq X \leq 70 + c) = P\left(\frac{-c}{3} \leq Z \leq \frac{c}{3}\right) = 2\Phi\left(\frac{c}{3}\right) - 1 = .95 \Rightarrow \Phi\left(\frac{c}{3}\right) = .9750$

$$\frac{c}{3} = 1.96 \Rightarrow c = 5.88$$

c. $10 \cdot P(\text{a single one is acceptable}) = 9.05$

d. $p = P(X < 73.84) = P(Z < 1.28) = .9$, so $P(Y \leq 8) = B(8; 10, .9) = .264$

43. The stated condition implies that 99% of the area under the normal curve with $\mu = 10$ and $\sigma = 2$ is to the left of $c - 1$, so $c - 1$ is the 99th percentile of the distribution. Thus $c - 1 = \mu + \sigma(2.33) = 20.155$, and $c = 21.155$.

44.

a. By symmetry, $P(-1.72 \leq Z \leq -.55) = P(.55 \leq Z \leq 1.72) = \Phi(1.72) - \Phi(.55)$

b. $P(-1.72 \leq Z \leq .55) = \Phi(.55) - \Phi(-1.72) = \Phi(.55) - [1 - \Phi(1.72)]$
No, symmetry of the Z curve about 0.

45. $X \sim N(3432, 482)$

a. $P(x > 4000) = P\left(Z > \frac{4000 - 3432}{482}\right) = P(z > 1.18)$
 $= 1 - \Phi(1.18) = 1 - .8810 = .1190$

$$P(3000 < x < 4000) = P\left(\frac{3000 - 3432}{482} < Z < \frac{4000 - 3432}{482}\right)$$

$$= \Phi(1.18) - \Phi(-.90) = .8810 - .1841 = .6969$$

b. $P(x < 2000 \text{ or } x > 5000) = P\left(Z < \frac{2000 - 3432}{482}\right) + P\left(Z > \frac{5000 - 3432}{482}\right)$
 $= \Phi(-2.97) + [1 - \Phi(3.25)] = .0015 + .0006 = .0021$

Chapter 4: Continuous Random Variables and Probability Distributions

- c. We will use the conversion 1 lb = 454 g, then 7 lbs = 3178 grams, and we wish to find

$$P(x > 3178) = P\left(Z > \frac{3178 - 3432}{482}\right) = 1 - \Phi(-.53) = .7019$$

- d. We need the top .0005 and the bottom .0005 of the distribution. Using the Z table, both .9995 and .0005 have multiple z values, so we will use a middle value, ± 3.295 . Then $3432 \pm (482)3.295 = 1844$ and 5020 , or the most extreme .1% of all birth weights are less than 1844 g and more than 5020 g.

- e. Converting to lbs yields mean 7.5595 and s.d. 1.0608. Then

$$P(x > 7) = P\left(Z > \frac{7 - 7.5595}{1.0608}\right) = 1 - \Phi(-.53) = .7019 \quad \text{This yields the same answer as in part c.}$$

46. We use a Normal approximation to the Binomial distribution: $X \sim b(x; 1000, .03) \approx N(30, 5.394)$

$$\begin{aligned} \text{a. } P(x \geq 40) &= 1 - P(x \leq 39) = 1 - P\left(Z \leq \frac{39.5 - 30}{5.394}\right) \\ &= 1 - \Phi(1.76) = 1 - .9608 = .0392 \end{aligned}$$

$$\text{b. } 5\% \text{ of } 1000 = 50: P(x \leq 50) = P\left(Z \leq \frac{50.5 - 30}{5.394}\right) = \Phi(3.80) \approx 1.00$$

47. $P(|X - \mu| \geq \sigma) = P(X \leq \mu - \sigma \text{ or } X \geq \mu + \sigma)$
 $= 1 - P(\mu - \sigma \leq X \leq \mu + \sigma) = 1 - P(-1 \leq Z \leq 1) = .3174$
 Similarly, $P(|X - \mu| \geq 2\sigma) = 1 - P(-2 \leq Z \leq 2) = .0456$
 And $P(|X - \mu| \geq 3\sigma) = 1 - P(-3 \leq Z \leq 3) = .0026$

- 48.

- a. $P(20 - .5 \leq X \leq 30 + .5) = P(19.5 \leq X \leq 30.5) = P(-1.1 \leq Z \leq 1.1) = .7286$
- b. $P(\text{at most } 30) = P(X \leq 30 + .5) = P(Z \leq 1.1) = .8643$
 $P(\text{less than } 30) = P(X < 30 - .5) = P(Z < .9) = .8159$

Chapter 4: Continuous Random Variables and Probability Distributions

49. P: .5 .6 .8
 μ : 12.5 15 20
 σ : 2.50 2.45 2.00
a.

	P(15 ≤ X ≤ 20)	P(14.5 ≤ normal ≤ 20.5)
.5	.212	P(.80 ≤ Z ≤ 3.20) = .2112
.6	.577	P(-.20 ≤ Z ≤ 2.24) = .5668
.8	.573	P(-2.75 ≤ Z ≤ .25) = .5957

b.

P(X ≤ 15)	P(normal ≤ 15.5)
.885	P(Z ≤ 1.20) = .8849
.575	P(Z ≤ .20) = .5793
.017	P(Z ≤ -2.25) = .0122

c.

P(20 ≤ X)	P(19.5 ≤ normal)
.002	.0026
.029	.0329
.617	.5987

50. P = .10; n = 200; np = 20, npq = 18

a. $P(X \leq 30) = \Phi\left(\frac{30 + .5 - 20}{\sqrt{18}}\right) = \Phi(2.47) = .9932$

b. $P(X < 30) = P(X \leq 29) = \Phi\left(\frac{29 + .5 - 20}{\sqrt{18}}\right) = \Phi(2.24) = .9875$

c. $P(15 \leq X \leq 25) = P(X \leq 25) - P(X \leq 14) = \Phi\left(\frac{25 + .5 - 20}{\sqrt{18}}\right) - \Phi\left(\frac{14 + .5 - 20}{\sqrt{18}}\right)$
 $\Phi(1.30) - \Phi(-1.30) = .9032 - .0968 = .8064$

51. N = 500, p = .4, $\mu = 200$, $\sigma = 10.9545$

a. $P(180 \leq X \leq 230) = P(179.5 \leq \text{normal} \leq 230.5) = P(-1.87 \leq Z \leq 2.78) = .9666$

b. $P(X < 175) = P(X \leq 174) = P(\text{normal} \leq 174.5) = P(Z \leq -2.33) = .0099$

Chapter 4: Continuous Random Variables and Probability Distributions

52. $P(X \leq \mu + \sigma[(100p)\text{th percentile for std normal}])$

$$P\left(\frac{X - \mu}{\sigma} \leq [\dots]\right) = P(Z \leq [\dots]) = p \text{ as desired}$$

53.

a. $F_y(y) = P(Y \leq y) = P(aX + b \leq y) = P\left(X \leq \frac{(y-b)}{a}\right)$ (for $a > 0$).

Now differentiate with respect to y to obtain

$$f_y(y) = F_y'(y) = \frac{1}{\sqrt{2\pi}a\sigma} e^{-\frac{1}{2a^2\sigma^2}[y-(a\mu+b)]^2}$$

so Y is normal with mean $a\mu + b$
and variance $a^2\sigma^2$.

b. Normal, mean $\frac{9}{5}(115) + 32 = 239$, variance = 12.96

54.

a. $P(Z \geq 1) \approx .5 \cdot \exp\left(\frac{83 + 351 + 562}{703 + 165}\right) = .1587$

b. $P(Z > 3) \approx .5 \cdot \exp\left(\frac{-2362}{399.3333}\right) = .0013$

c. $P(Z > 4) \approx .5 \cdot \exp\left(\frac{-3294}{340.75}\right) = .0000317$, so
 $P(-4 < Z < 4) \approx 1 - 2(.0000317) = .999937$

d. $P(Z > 5) \approx .5 \cdot \exp\left(\frac{-4392}{305.6}\right) = .00000029$

Section 4.4

55.

a. $\Gamma(6) = 5! = 120$

b. $\Gamma\left(\frac{5}{2}\right) = \frac{3}{2}\Gamma\left(\frac{1}{2}\right) = \frac{3}{2} \cdot \frac{1}{2} \cdot \Gamma\left(\frac{1}{2}\right) = \left(\frac{3}{4}\right)\sqrt{\pi} \approx 1.329$

c. $F(4;5) = .371$ from row 4, column 5 of Table A.4

d. $F(5;4) = .735$

e. $F(0;4) = P(X \leq 0; \alpha = 4) = 0$

56.

a. $P(X \leq 5) = F(5;7) = .238$

b. $P(X < 5) = P(X \leq 5) = .238$

c. $P(X > 8) = 1 - P(X \leq 8) = 1 - F(8;7) = .313$

d. $P(3 \leq X \leq 8) = F(8;7) - F(3;7) = .653$

e. $P(3 < X < 8) = .653$

f. $P(X < 4 \text{ or } X > 6) = 1 - P(4 \leq X \leq 6) = 1 - [F(6;7) - F(4;7)] = .713$

57.

a. $\mu = 20, \sigma^2 = 80 \Rightarrow \alpha\beta = 20, \alpha\beta^2 = 80 \Rightarrow \beta = \frac{80}{20}, \alpha = 5$

b. $P(X \leq 24) = F\left(\frac{24}{4}; 5\right) = F(6;5) = .715$

c. $P(20 \leq X \leq 40) = F(10;5) - F(5;5) = .411$

58.

$\mu = 24, \sigma^2 = 144 \Rightarrow \alpha\beta = 24, \alpha\beta^2 = 144 \Rightarrow \beta = 6, \alpha = 4$

a. $P(12 \leq X \leq 24) = F(4;4) - F(2;4) = .424$

b. $P(X \leq 24) = F(4;4) = .567$, so while the mean is 24, the median is less than 24. ($P(X \leq \tilde{\mu}) = .5$); This is a result of the positive skew of the gamma distribution.

Chapter 4: Continuous Random Variables and Probability Distributions

- c. We want a value of X for which $F(X;4)=.99$. In table A.4, we see $F(10;4)=.990$. So with $\beta = 6$, the 99th percentile = $6(10)=60$.
- d. We want a value of X for which $F(X;4)=.995$. In the table, $F(11;4)=.995$, so $t = 6(11)=66$. At 66 weeks, only .5% of all transistors would still be operating.

59.

a. $E(X) = \frac{1}{\lambda} = 1$

b. $\sigma = \frac{1}{\lambda} = 1$

c. $P(X \leq 4) = 1 - e^{-(1)(4)} = 1 - e^{-4} = .982$

d. $P(2 \leq X \leq 5) = 1 - e^{-(1)(5)} - [1 - e^{-(1)(2)}] = e^{-2} - e^{-5} = .129$

60.

a. $P(X \leq 100) = 1 - e^{-(100)(.01386)} = 1 - e^{-1.386} = .7499$
 $P(X \leq 200) = 1 - e^{-(200)(.01386)} = 1 - e^{-2.772} = .9375$
 $P(100 \leq X \leq 200) = P(X \leq 200) - P(X \leq 100) = .9375 - .7499 = .1876$

b. $\mu = \frac{1}{.01386} = 72.15, \sigma = 72.15$
 $P(X > \mu + 2\sigma) = P(X > 72.15 + 2(72.15)) = P(X > 216.45) =$
 $1 - [1 - e^{-(216.45)(.01386)}] = e^{-2.9999} = .0498$

c. $.5 = P(X \leq \tilde{\mu}) \Rightarrow 1 - e^{-(\tilde{\mu})(.01386)} = .5 \Rightarrow e^{-(\tilde{\mu})(.01386)} = .5$
 $-\tilde{\mu}(.01386) = \ln(.5) = .693 \Rightarrow \tilde{\mu} = 50$

61. Mean = $\frac{1}{\lambda} = 25,000$ implies $\lambda = .00004$

a. $P(X > 20,000) = 1 - P(X \leq 20,000) = 1 - F(20,000; .00004) = e^{-(.00004)(20,000)} = .449$
 $P(X \leq 30,000) = F(30,000; .00004) = e^{-1.2} = .699$
 $P(20,000 \leq X \leq 30,000) = .699 - .551 = .148$

b. $\sigma = \frac{1}{\lambda} = 25,000$, so $P(X > \mu + 2\sigma) = P(x > 75,000) =$
 $1 - F(75,000; .00004) = .05.$
 Similarly, $P(X > \mu + 3\sigma) = P(x > 100,000) = .018$

62.

- a. $E(X) = \alpha\beta = n \frac{1}{\lambda} = \frac{n}{\lambda}$; for $\lambda = .5$, $n = 10$, $E(X) = 20$
- b. $P(X \leq 30) = F\left(\frac{30}{2}; 10\right) = F(15; 10) = .930$
- c. $P(X \leq t) = P(\text{at least } n \text{ events in time } t) = P(Y \geq n)$ when $Y \sim \text{Poisson}$ with parameter λt .
Thus $P(X \leq t) = 1 - P(Y < n) = 1 - P(Y \leq n - 1) = 1 - \sum_{k=0}^{n-1} \frac{e^{-\lambda t} (\lambda t)^k}{k!}$.

63.

- a. $\{X \geq t\} = A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5$
- b. $P(X \geq t) = P(A_1) \cdot P(A_2) \cdot P(A_3) \cdot P(A_4) \cdot P(A_5) = (e^{-\lambda t})^5 = e^{-.05t}$, so $F_X(t) = P(X \leq t) = 1 - e^{-.05t}$, $f_X(t) = .05e^{-.05t}$ for $t \geq 0$. Thus X also has an exponential distribution, but with parameter $\lambda = .05$.
- c. By the same reasoning, $P(X \leq t) = 1 - e^{-n\lambda t}$, so X has an exponential distribution with parameter $n\lambda$.

64.

With $x_p = (100p)\text{th percentile}$, $p = F(x_p) = 1 - e^{-\lambda x_p} \Rightarrow e^{-\lambda x_p} = 1 - p$,
 $\Rightarrow -\lambda x_p = \ln(1 - p) \Rightarrow x_p = \frac{-[\ln(1 - p)]}{\lambda}$. For $p = .5$, $x_{.5} = \tilde{\mu} = \frac{.693}{\lambda}$.

65.

- a. $\{X^2 \leq y\} = \{-\sqrt{y} \leq X \leq \sqrt{y}\}$
- b. $P(X^2 \leq y) = \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz$. Now differentiate with respect to y to obtain the chi-squared p.d.f. with $v = 1$.

Section 4.5

66.

$$\text{a. } E(X) = 3\Gamma\left(1 + \frac{1}{2}\right) = 3 \cdot \frac{1}{2} \cdot \Gamma\left(\frac{1}{2}\right) = 2.66,$$

$$\text{Var}(X) = 9\left[\Gamma(1+1) - \Gamma^2\left(1 + \frac{1}{2}\right)\right] = 1.926$$

$$\text{b. } P(X \leq 6) = 1 - e^{-(6/3)^\alpha} = 1 - e^{-(6/3)^2} = 1 - e^{-4} = .982$$

$$\text{c. } P(1.5 \leq X \leq 6) = 1 - e^{-(6/3)^2} - \left[1 - e^{-(1.5/3)^2}\right] = e^{-.25} - e^{-4} = .760$$

67.

$$\text{a. } P(X \leq 250) = F(250; 2.5, 200) = 1 - e^{-(250/200)^{2.5}} = 1 - e^{-1.75} \approx .8257$$

$$P(X < 250) = P(X \leq 250) \approx .8257$$

$$P(X > 300) = 1 - F(300; 2.5, 200) = e^{-(1.5)^{2.5}} = .0636$$

$$\text{b. } P(100 \leq X \leq 250) = F(250; 2.5, 200) - F(100; 2.5, 200) \approx .8257 - .162 = .6637$$

c. The median $\tilde{\mu}$ is requested. The equation $F(\tilde{\mu}) = .5$ reduces to

$$.5 = e^{-(\tilde{\mu}/200)^{2.5}}, \text{ i.e., } \ln(.5) \approx -\left(\frac{\tilde{\mu}}{200}\right)^{2.5}, \text{ so } \tilde{\mu} = (.6931)^4(200) = 172.727.$$

68.

$$\text{a. } \text{For } x > 3.5, F(x) = P(X \leq x) = P(X - 3.5 \leq x - 3.5) = 1 - e^{-\left[\frac{(x-3.5)}{1.5}\right]^2}$$

$$\text{b. } E(X - 3.5) = 1.5\Gamma\left(\frac{3}{2}\right) = 1.329 \text{ so } E(X) = 4.829$$

$$\text{Var}(X) = \text{Var}(X - 3.5) = (1.5)^2 \left[\Gamma(2) - \Gamma^2\left(\frac{3}{2}\right) \right] = .483$$

$$\text{c. } P(X > 5) = 1 - P(X \leq 5) = 1 - \left[1 - e^{-1}\right] = e^{-1} = .368$$

$$\text{d. } P(5 \leq X \leq 8) = 1 - e^{-9} - \left[1 - e^{-1}\right] = e^{-1} - e^{-9} = .3679 - .0001 = .3678$$

69. $\mu = \int_0^{\infty} x \cdot \frac{\alpha}{\beta^\alpha} x^{\alpha-1} e^{-\left(\frac{x}{\beta}\right)^\alpha} dx = (\text{after } y = \left(\frac{x}{\beta}\right)^\alpha, dy = \frac{\alpha x^{\alpha-1}}{\beta^\alpha} dx)$

$$\beta \int_0^{\infty} y^{1/\alpha} e^{-y} dy = \beta \cdot \Gamma\left(1 + \frac{1}{\alpha}\right) \text{ by definition of the gamma function.}$$

70.

a. $.5 = F(\tilde{\mu}) = 1 - e^{-(\mu/3)^2} \Rightarrow$

$$e^{-\mu^2/9} = .5 \Rightarrow \tilde{\mu}^2 = -9 \ln(.5) = 6.2383 \Rightarrow \tilde{\mu} = 2.50$$

b. $1 - e^{-[(\tilde{\mu}-3.5)/1.5]^2} = .5 \Rightarrow (\tilde{\mu} - 3.5)^2 = -2.25 \ln(.5) = 1.5596 \Rightarrow \tilde{\mu} = 4.75$

c. $P = F(x_p) = 1 - e^{-\left(\frac{x_p}{\beta}\right)^\alpha} \Rightarrow (x_p/\beta)^\alpha = -\ln(1-p) \Rightarrow x_p = \beta[-\ln(1-p)]^{1/\alpha}$

d. The desired value of t is the 90th percentile (since 90% will not be refused and 10% will be). From c, the 90th percentile of the distribution of $X - 3.5$ is $1.5[-\ln(.1)]^{1/2} = 2.27661$, so $t = 3.5 + 2.2761 = 5.7761$

71. $X \sim \text{Weibull: } \alpha=20, \beta=100$

a. $F(x, 20, \beta) = 1 - e^{-\left(\frac{x}{\beta}\right)^\alpha} = 1 - e^{-\left(\frac{105}{100}\right)^{20}} = 1 - .070 = .930$

b. $F(105) - F(100) = .930 - (1 - e^{-1}) = .930 - .632 = .298$

c. $.50 = 1 - e^{-\left(\frac{x}{100}\right)^{20}} \Rightarrow e^{-\left(\frac{x}{100}\right)^{20}} = .50 \Rightarrow -\left(\frac{x}{100}\right)^{20} = \ln(.50)$
 $\left(\frac{-x}{100}\right) = \sqrt[20]{\ln(.50)} \Rightarrow -x = 100 \left(\sqrt[20]{\ln(.50)}\right) \Rightarrow x = 98.18$

72.

a. $E(X) = e^{\left(\mu + \frac{\sigma^2}{2}\right)} = e^{4.82} = 123.97$

$$V(X) = \left(e^{(2(4.5) + .8^2)}\right) \cdot (e^{-.8} - 1) = (15,367.34)(.8964) = 13,776.53$$

$$\sigma = 117.373$$

b. $P(x \leq 100) = P\left(z \leq \frac{\ln(100) - 4.5}{.8}\right) = \Phi(0.13) = .5517$

c. $P(x \geq 200) = P\left(z \geq \frac{\ln(200) - 4.5}{.8}\right) = 1 - \Phi(1.00) = 1 - .8413 = .1587 = P(x > 200)$

Chapter 4: Continuous Random Variables and Probability Distributions

73.

a. $E(X) = e^{3.5+(1.2)^2/2} = 68.0335$; $V(X) = e^{2(3.5)+(1.2)^2} \cdot (e^{(1.2)^2} - 1) = 14907.168$;
 $\sigma_x = 122.0949$

b. $P(50 \leq X \leq 250) = P\left(z \leq \frac{\ln(250) - 3.5}{1.2}\right) - P\left(z \leq \frac{\ln(50) - 3.5}{1.2}\right)$
 $P(Z \leq 1.68) - P(Z \leq .34) = .9535 - .6331 = .3204$.

c. $P(X \leq 68.0335) = P\left(z \leq \frac{\ln(68.0335) - 3.5}{1.2}\right) = P(Z \leq .60) = .7257$. The lognormal distribution is not a symmetric distribution.

74.

a. $.5 = F(\tilde{\mu}) = \Phi\left(\frac{\ln(\tilde{\mu}) - \mu}{\sigma}\right)$, (where $\tilde{\mu}$ refers to the lognormal distribution and μ and σ to the normal distribution). Since the median of the standard normal distribution is 0, $\frac{\ln(\tilde{\mu}) - \mu}{\sigma} = 0$, so $\ln(\tilde{\mu}) = \mu \Rightarrow \tilde{\mu} = e^\mu$. For the power distribution, $\tilde{\mu} = e^{3.5} = 33.12$

b. $1 - \alpha = \Phi(z_\alpha) = P(Z \leq z_\alpha) = P\left(\frac{\ln(X) - \mu}{\sigma} \leq z_\alpha\right) = P(\ln(X) \leq \mu + \sigma z_\alpha)$
 $= P(X \leq e^{\mu + \sigma z_\alpha})$, so the $100(1 - \alpha)$ th percentile is $e^{\mu + \sigma z_\alpha}$. For the power distribution, the 95th percentile is $e^{3.5+(1.645)(1.2)} = e^{5.474} = 238.41$

75.

a. $E(X) = e^{5+(.01)/2} = e^{5.005} = 149.157$; $\text{Var}(X) = e^{10+(.01)} \cdot (e^{.01} - 1) = 223.594$

b. $P(X > 125) = 1 - P(X \leq 125) =$
 $= 1 - P\left(z \leq \frac{\ln(125) - 5}{.1}\right) = 1 - \Phi(-1.72) = .9573$

c. $P(110 \leq X \leq 125) = \Phi(-1.72) - \Phi\left(\frac{\ln(110) - 5}{.1}\right) = .0427 - .0013 = .0414$

d. $\tilde{\mu} = e^5 = 148.41$ (continued)

e. $P(\text{any particular one has } X > 125) = .9573 \Rightarrow \text{expected \#} = 10(.9573) = 9.573$

f. We wish the 5th percentile, which is $e^{5+(-1.645)(.1)} = 125.90$

76.

a. $E(X) = e^{1.9+9^2/2} = 10.024$; $\text{Var}(X) = e^{3.8+(.81)} \cdot (e^{.81} - 1) = 125.395$, $\sigma_x = 11.20$

b. $P(X \leq 10) = P(\ln(X) \leq 2.3026) = P(Z \leq .45) = .6736$
 $P(5 \leq X \leq 10) = P(1.6094 \leq \ln(X) \leq 2.3026)$
 $= P(-.32 \leq Z \leq .45) = .6736 - .3745 = .2991$

77.

The point of symmetry must be $\frac{1}{2}$, so we require that $f\left(\frac{1}{2} - \mu\right) = f\left(\frac{1}{2} + \mu\right)$, i.e.,
 $\left(\frac{1}{2} - \mu\right)^{\alpha-1} \left(\frac{1}{2} + \mu\right)^{\beta-1} = \left(\frac{1}{2} + \mu\right)^{\alpha-1} \left(\frac{1}{2} - \mu\right)^{\beta-1}$, which in turn implies that $\alpha = \beta$.

78.

a. $E(X) = \frac{5}{(5+2)} = \frac{5}{7} = .714$, $V(X) = \frac{10}{(49)(8)} = .0255$

b. $f(x) = \frac{\Gamma(7)}{\Gamma(5)\Gamma(2)} \cdot x^4 \cdot (1-x) = 30(x^4 - x^5)$ for $0 \leq X \leq 1$,
 so $P(X \leq .2) = \int_0^{.2} 30(x^4 - x^5) dx = .0016$

c. $P(.2 \leq X \leq .4) = \int_{.2}^{.4} 30(x^4 - x^5) dx = .03936$

d. $E(1-X) = 1 - E(X) = 1 - \frac{5}{7} = \frac{2}{7} = .286$

79.

a. $E(X) = \int_0^1 x \cdot \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1} dx = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 x^\alpha (1-x)^{\beta-1} dx$
 $\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot \frac{\Gamma(\alpha+1)\Gamma(\beta)}{\Gamma(\alpha+\beta+1)} = \frac{\alpha\Gamma(\alpha)}{\Gamma(\alpha)\Gamma(\beta)} \cdot \frac{\Gamma(\alpha+\beta)}{(\alpha+\beta)\Gamma(\alpha+\beta)} = \frac{\alpha}{\alpha+\beta}$

b. $E[(1-X)^m] = \int_0^1 (1-x)^m \cdot \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1} dx$
 $= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 x^{\alpha-1} (1-x)^{m+\beta-1} dx = \frac{\Gamma(\alpha+\beta) \cdot \Gamma(m+\beta)}{\Gamma(\alpha+\beta+m)\Gamma(\beta)}$
 For $m = 1$, $E(1-X) = \frac{\beta}{\alpha+\beta}$.

80.

$$\text{a. } E(Y) = 10 \Rightarrow E\left(\frac{Y}{20}\right) = \frac{1}{2} = \frac{\alpha}{\alpha + \beta}; \text{Var}(Y) = \frac{100}{7} \Rightarrow \text{Var}\left(\frac{Y}{20}\right) = \frac{100}{2800} = \frac{1}{28}$$

$$\frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)} \Rightarrow \alpha = 3, \beta = 3, \text{ after some algebra.}$$

$$\text{b. } P(8 \leq X \leq 12) = F\left(\frac{12}{20}; 3, 3\right) - F\left(\frac{8}{20}; 3, 3\right) = F(.6; 3, 3) - F(.4; 3, 3).$$

The standard density function here is $30y^2(1 - y)^2$,

$$\text{so } P(8 \leq X \leq 12) = \int_{.4}^{.6} 30y^2(1 - y)^2 dy = .365.$$

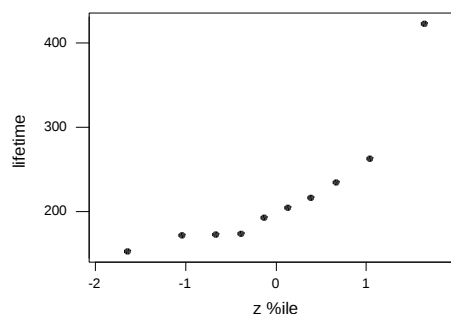
$$\text{c. } \text{We expect it to snap at 10, so } P(Y < 8 \text{ or } Y > 12) = 1 - P(8 \leq X \leq 12) \\ = 1 - .365 = .665.$$

Section 4.6

81. The given probability plot is quite linear, and thus it is quite plausible that the tension distribution is normal.

82. The z percentiles and observations are as follows:

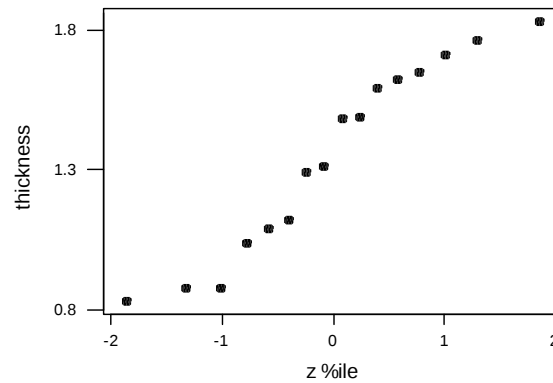
percentile	observation
-1.645	152.7
-1.040	172.0
-0.670	172.5
-0.390	173.3
-0.130	193.0
0.130	204.7
0.390	216.5
0.670	234.9
1.040	262.6
1.645	422.6



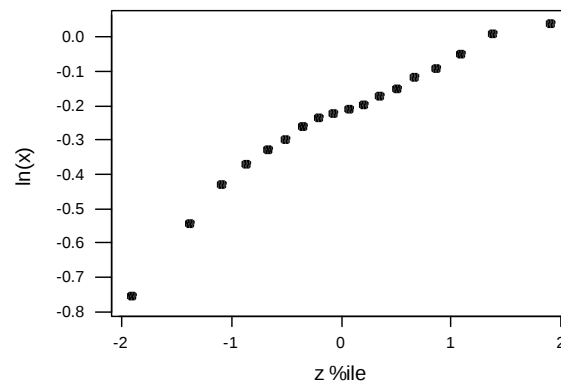
The accompanying plot is quite straight except for the point corresponding to the largest observation. This observation is clearly much larger than what would be expected in a normal random sample. Because of this outlier, it would be inadvisable to analyze the data using any inferential method that depended on assuming a normal population distribution.

Chapter 4: Continuous Random Variables and Probability Distributions

83. The z percentile values are as follows: -1.86, -1.32, -1.01, -0.78, -0.58, -0.40, -0.24, -0.08, 0.08, 0.24, 0.40, 0.58, 0.78, 1.01, 1.30, and 1.86. The accompanying probability plot is reasonably straight, and thus it would be reasonable to use estimating methods that assume a normal population distribution.

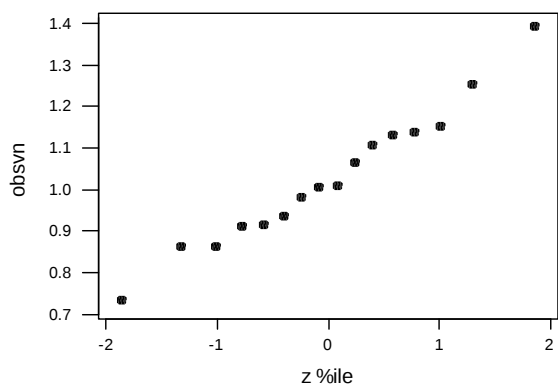


84. The Weibull plot uses $\ln(\text{observations})$ and the z percentiles of the p_i values given. The accompanying probability plot appears sufficiently straight to lead us to agree with the argument that the distribution of fracture toughness in concrete specimens could well be modeled by a Weibull distribution.

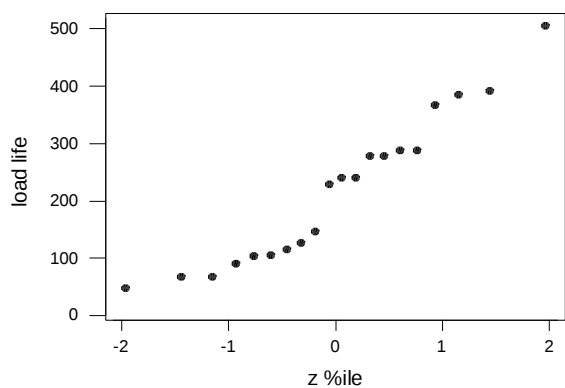


Chapter 4: Continuous Random Variables and Probability Distributions

85. The (z percentile, observation) pairs are (-1.66, .736), (-1.32, .863), (-1.01, .865), (-.78, .913), (-.58, .915), (-.40, .937), (-.24, .983), (-.08, 1.007), (.08, 1.011), (.24, 1.064), (.40, 1.109), (.58, 1.132), (.78, 1.140), (1.01, 1.153), (1.32, 1.253), (1.86, 1.394). The accompanying probability plot is very straight, suggesting that an assumption of population normality is extremely plausible.

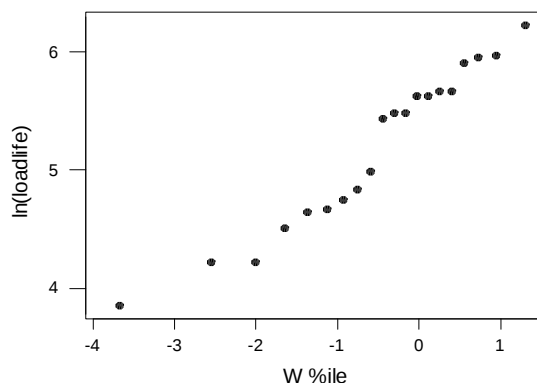


86. a. The 10 largest z percentiles are 1.96, 1.44, 1.15, .93, .76, .60, .45, .32, .19 and .06; the remaining 10 are the negatives of these values. The accompanying normal probability plot is reasonably straight. An assumption of population distribution normality is plausible.

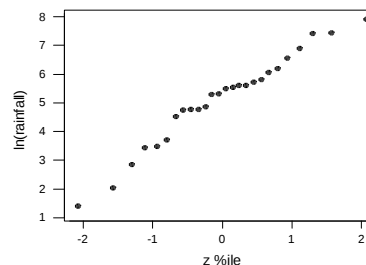
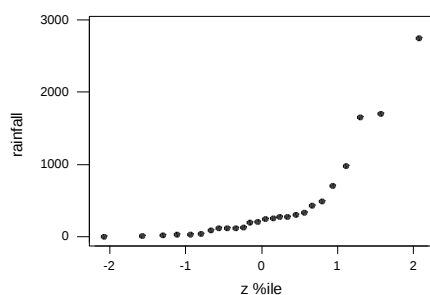


Chapter 4: Continuous Random Variables and Probability Distributions

- b. For a Weibull probability plot, the natural logs of the observations are plotted against extreme value percentiles; these percentiles are -3.68, -2.55, -2.01, -1.65, -1.37, -1.13, -.93, -.76, -.59, -.44, -.30, -.16, -.02, .12, .26, .40, .56, .73, .95, and 1.31. The accompanying probability plot is roughly as straight as the one for checking normality (a plot of $\ln(x)$ versus the z percentiles, appropriate for checking the plausibility of a lognormal distribution, is also reasonably straight - any of 3 different families of population distributions seems plausible.)

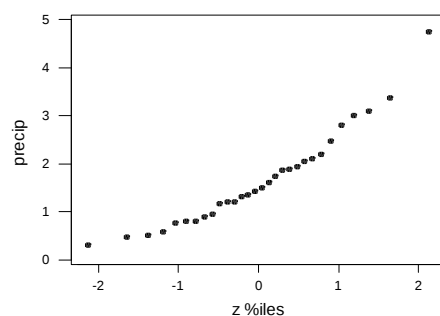


87. To check for plausibility of a lognormal population distribution for the rainfall data of Exercise 81 in Chapter 1, take the natural logs and construct a normal probability plot. This plot and a normal probability plot for the original data appear below. Clearly the log transformation gives quite a straight plot, so lognormality is plausible. The curvature in the plot for the original data implies a positively skewed population distribution - like the lognormal distribution.

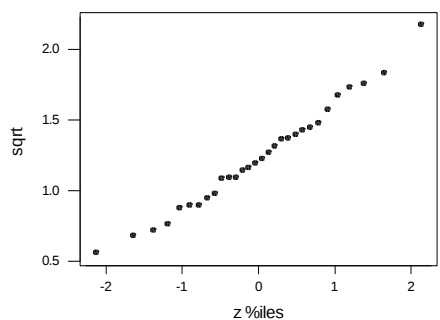


88.

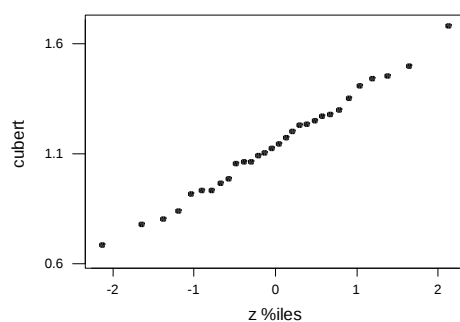
- a. The plot of the original (untransformed) data appears somewhat curved.



- b. The square root transformation results in a very straight plot. It is reasonable that this distribution is normally distributed.



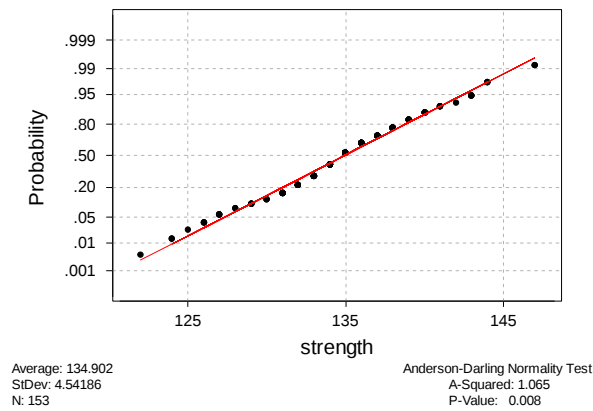
- c. The cube root transformation also results in a very straight plot. It is very reasonable that the distribution is normally distributed.



Chapter 4: Continuous Random Variables and Probability Distributions

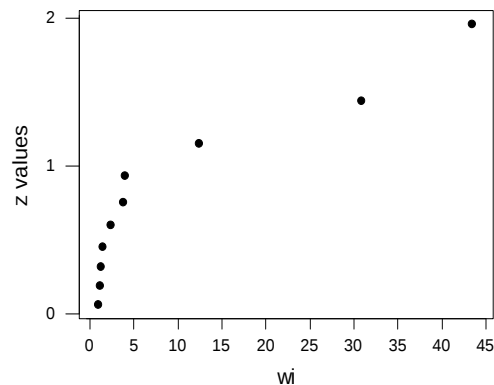
89. The pattern in the plot (below, generated by Minitab) is quite linear. It is very plausible that strength is normally distributed.

Normal Probability Plot



90. We use the data (table below) to create the desired plot.

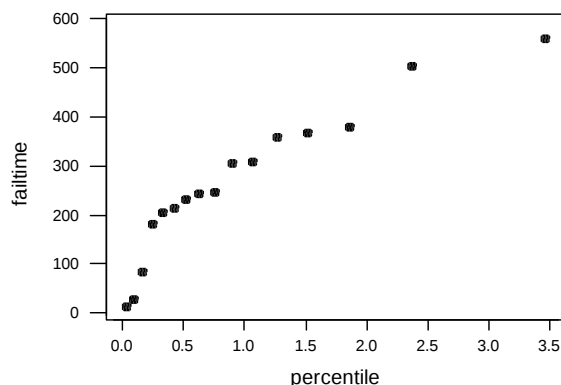
ordered absolute values (w's)	probabilities	z values
0.89	0.525	0.063
1.15	0.575	0.19
1.27	0.625	0.32
1.44	0.675	0.454
2.34	0.725	0.6
3.78	0.775	0.755
3.96	0.825	0.935
12.38	0.875	1.15
30.84	0.925	1.44
43.4	0.975	1.96



This half-normal plot reveals some extreme values, without which the distribution may appear to be normal.

Chapter 4: Continuous Random Variables and Probability Distributions

91. The $(100p)^{\text{th}}$ percentile $\eta(p)$ for the exponential distribution with $\lambda = 1$ satisfies $F(\eta(p)) = 1 - \exp[-\eta(p)] = p$, i.e., $\eta(p) = -\ln(1 - p)$. With $n = 16$, we need $\eta(p)$ for $p = \frac{5}{16}, \frac{15}{16}, \dots, \frac{15.5}{16}$. These are .032, .398, .170, .247, .330, .421, .521, .633, .758, .901, 1.068, 1.269, 1.520, 1.856, 2.367, 3.466. this plot exhibits substantial curvature, casting doubt on the assumption of an exponential population distribution. Because λ is a scale parameter (as is σ for the normal family), $\lambda = 1$ can be used to assess the plausibility of the entire exponential family.



Supplementary Exercises

92.

- a. $P(10 \leq X \leq 20) = \frac{10}{25} = .4$
- b. $P(X \geq 10) = P(10 \leq X \leq 25) = \frac{15}{25} = .6$
- c. For $0 \leq X \leq 25$, $F(x) = \int_0^x \frac{1}{25} dy = \frac{x}{25}$. $F(x) = 0$ for $x < 0$ and $= 1$ for $x > 25$.
- d. $E(X) = \frac{(A + B)}{2} = \frac{(0 + 25)}{2} = 12.5$; $\text{Var}(X) = \frac{(B - A)^2}{12} = \frac{625}{12} = 52.083$

93.

a. For $0 \leq Y \leq 25$, $F(y) = \frac{1}{24} \int_0^y \left(u - \frac{u^2}{12}\right) du = \frac{1}{24} \left(\frac{u^2}{2} - \frac{u^3}{36}\right) \Big|_0^y$. Thus

$$F(y) = \begin{cases} 0 & y < 0 \\ \frac{1}{48} \left(y^2 - \frac{y^3}{18}\right) & 0 \leq y \leq 12 \\ 1 & y > 12 \end{cases}$$

b. $P(Y \leq 4) = F(4) = .259$, $P(Y > 6) = 1 - F(6) = .5$
 $P(4 \leq X \leq 6) = F(6) - F(4) = .5 - .259 = .241$

c. $E(Y) = \frac{1}{24} \int_0^{12} y^2 \left(1 - \frac{y}{12}\right) dy = \frac{1}{24} \left[\frac{y^3}{3} - \frac{y^4}{48}\right]_0^{12} = 6$
 $E(Y^2) = \frac{1}{24} \int_0^{12} y^3 \left(1 - \frac{y}{12}\right) dy = 43.2$, so $V(Y) = 43.2 - 36 = 7.2$

d. $P(Y < 4 \text{ or } Y > 8) = 1 - P(4 \leq Y \leq 8) = .518$

e. the shorter segment has length $\min(Y, 12 - Y)$ so
 $E[\min(Y, 12 - Y)] = \int_0^{12} \min(y, 12 - y) \cdot f(y) dy = \int_0^6 \min(y, 12 - y) \cdot f(y) dy$
 $+ \int_6^{12} \min(y, 12 - y) \cdot f(y) dy = \int_0^6 y \cdot f(y) dy + \int_6^{12} (12 - y) \cdot f(y) dy = \frac{90}{24} = 3.75$

94.

a. Clearly $f(x) \geq 0$. The c.d.f. is, for $x > 0$,

$$F(x) = \int_{-\infty}^x f(y) dy = \int_0^x \frac{32}{(y+4)^3} dy = -\frac{1}{2} \cdot \frac{32}{(y+4)^2} \Big|_0^x = 1 - \frac{16}{(x+4)^2}$$

($F(x) = 0$ for $x \leq 0$.)

Since $F(\infty) = \int_{-\infty}^{\infty} f(y) dy = 1$, $f(x)$ is a legitimate pdf.

b. See above

c. $P(2 \leq X \leq 5) = F(5) - F(2) = 1 - \frac{16}{81} - \left(1 - \frac{16}{36}\right) = .247$

(continued)

Chapter 4: Continuous Random Variables and Probability Distributions

$$\text{d. } E(x) = \int_{-\infty}^{\infty} x \cdot f(x) dx = \int_{-\infty}^{\infty} x \cdot \frac{32}{(x+4)^3} dx = \int_0^{\infty} (x+4-4) \cdot \frac{32}{(x+4)^3} dx$$

$$= \int_0^{\infty} \frac{32}{(x+4)^2} dx - 4 \int_0^{\infty} \frac{32}{(x+4)^3} dx = 8 - 4 = 4$$

$$\text{e. } E(\text{salvage value}) = \int_0^{\infty} \frac{100}{x+4} \cdot \frac{32}{(y+4)^3} dx = 3200 \int_0^{\infty} \frac{1}{(y+4)^4} dy = \frac{3200}{(3)(64)} = 16.67$$

95.

a. By differentiation,

$$f(x) = \begin{cases} x^2 & 0 \leq x < 1 \\ \frac{7}{4} - \frac{3}{4}x & 1 \leq x \leq \frac{7}{3} \\ 0 & \text{otherwise} \end{cases}$$

$$\text{b. } P(.5 \leq X \leq 2) = F(2) - F(.5) = 1 - \frac{1}{2} \left(\frac{7}{3} - 2 \right) \left(\frac{7}{4} - \frac{3}{4} \cdot 2 \right) - \frac{(.5)^3}{3} = \frac{11}{12} = .917$$

$$\text{c. } E(X) = \int_0^1 x \cdot x^2 dx + \int_1^{7/3} x \cdot \left(\frac{7}{4} - \frac{3}{4}x \right) dx = \frac{131}{108} = 1.213$$

96. $\mu = 40 \text{ V}; \sigma = 1.5 \text{ V}$

$$\text{a. } P(39 < X < 42) = \Phi\left(\frac{42-40}{1.5}\right) - \Phi\left(\frac{39-40}{1.5}\right) \\ = \Phi(1.33) - \Phi(-.67) = .9082 - .2514 = .6568$$

b. We desire the 85th percentile: $40 + (1.04)(1.5) = 41.56$

$$\text{c. } P(X > 42) = 1 - P(X \leq 42) = 1 - \Phi\left(\frac{42-40}{1.5}\right) = 1 - \Phi(1.33) = .0918$$

Let D represent the number of diodes out of 4 with voltage exceeding 42.

$$P(D \geq 1) = 1 - P(D = 0) = 1 - \binom{4}{0} (.0918)^0 (.9082)^4 = 1 - .6803 = .3197$$

Chapter 4: Continuous Random Variables and Probability Distributions

97. $\mu = 137.2 \text{ oz.}; \sigma = 1.6 \text{ oz}$

a. $P(X > 135) = 1 - \Phi\left(\frac{135 - 137.2}{1.6}\right) = 1 - \Phi(-1.38) = 1 - .0838 = .9162$

b. With $Y =$ the number among ten that contain more than 135 oz,
 $Y \sim \text{Bin}(10, .9162)$, so $P(Y \geq 8) = b(8; 10, .9162) + b(9; 10, .9162)$
 $+ b(10; 10, .9162) = .9549$.

c. $\mu = 137.2; \frac{135 - 137.2}{\sigma} = -1.65 \Rightarrow \sigma = 1.33$

98.

a. Let $S =$ defective. Then $p = P(S) = .05$; $n = 250 \Rightarrow \mu = np = 12.5, \sigma = 3.446$. The random variable $X =$ the number of defectives in the batch of 250. $X \sim \text{Binomial}$. Since $np = 12.5 \geq 10$, and $nq = 237.5 \geq 10$, we can use the normal approximation.

$$P(X_{\text{bin}} \geq 25) \approx 1 - \Phi\left(\frac{24.5 - 12.5}{3.446}\right) = 1 - \Phi(3.48) = 1 - .9997 = .0003$$

b. $P(X_{\text{bin}} = 10) \approx P(X_{\text{norm}} \leq 10.5) - P(X_{\text{norm}} \leq 9.5)$
 $= \Phi(-.58) - \Phi(-.87) = .2810 - .1922 = .0888$

99.

a. $P(X > 100) = 1 - \Phi\left(\frac{100 - 96}{14}\right) = 1 - \Phi(.29) = 1 - .6141 = .3859$

b. $P(50 < X < 80) = \Phi\left(\frac{80 - 96}{14}\right) - \Phi\left(\frac{50 - 96}{14}\right)$
 $= \Phi(-1.5) - \Phi(-3.29) = .1271 - .0005 = .1266$.

c. $a = 5^{\text{th}}$ percentile $= 96 + (-1.645)(14) = 72.97$.
 $b = 95^{\text{th}}$ percentile $= 96 + (1.645)(14) = 119.03$. The interval (72.97, 119.03) contains the central 90% of all grain sizes.

Chapter 4: Continuous Random Variables and Probability Distributions

100.

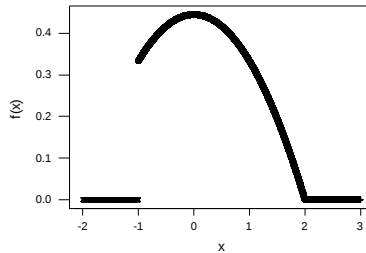
- a. $F(x) = 0$ for $x < 1$ and $= 1$ for $x > 3$. For $1 \leq x \leq 3$, $F(x) = \int_{-\infty}^x f(y)dy$

$$= \int_{-\infty}^1 0dy + \int_1^x \frac{3}{2} \cdot \frac{1}{y^2} dy = 1.5 \left(1 - \frac{1}{x} \right)$$
- b. $P(X \leq 2.5) = F(2.5) = 1.5(1 - .4) = .9$; $P(1.5 \leq x \leq 2.5) = F(2.5) - F(1.5) = .4$
- c. $E(X) = \int_1^3 x \cdot \frac{3}{2} \cdot \frac{1}{x^2} dx = \frac{3}{2} \int_1^3 \frac{1}{x} dx = 1.5 \ln(x) \Big|_1^3 = 1.648$
- d. $E(X^2) = \int_1^3 x^2 \cdot \frac{3}{2} \cdot \frac{1}{x^2} dx = \frac{3}{2} \int_1^3 dx = 3$, so $V(X) = E(X^2) - [E(X)]^2 = .284$,
 $\sigma = .553$
- e.
$$h(x) = \begin{cases} 0 & 1 \leq x \leq 1.5 \\ x - 1.5 & 1.5 \leq x \leq 2.5 \\ 1 & 2.5 \leq x \leq 3 \end{cases}$$

 so $E[h(X)] = \int_{1.5}^{2.5} (x - 1.5) \cdot \frac{3}{2} \cdot \frac{1}{x^2} dx + \int_{2.5}^3 1 \cdot \frac{3}{2} \cdot \frac{1}{x^2} dx = .267$

101.

a.



- b. $F(x) = 0$ for $x < -1$ or $= 1$ for $x > 2$. For $-1 \leq x \leq 2$,

$$F(x) = \int_{-1}^x \frac{1}{9} (4 - y^2) dy = \frac{1}{9} \left(4x - \frac{x^3}{3} \right) + \frac{11}{27}$$
- c. The median is 0 iff $F(0) = .5$. Since $F(0) = \frac{11}{27}$, this is not the case. Because $\frac{11}{27} < .5$, the median must be greater than 0.
- d. Y is a binomial r.v. with $n = 10$ and $p = P(X > 1) = 1 - F(1) = \frac{5}{27}$

102.

- a. $E(X) = \frac{1}{\lambda} = 1.075, \sigma = \frac{1}{\lambda} = 1.075$
- b. $P(3.0 < X) = 1 - P(X \leq 3.0) = 1 - F(3.0) = 3^{-.93(3.0)} = .0614$
 $P(1.0 \leq X \leq 3.0) = F(3.0) - F(1.0) = .333$
- c. The 90th percentile is requested; denoting it by c , we have
 $.9 = F(c) = 1 - e^{-(.93)c}$, whence $c = \frac{\ln(.1)}{(-.93)} = 2.476$

103.

- a. $P(X \leq 150) = \exp\left[-\exp\left(\frac{-(150-150)}{90}\right)\right] = \exp[-\exp(0)] = \exp(-1) = .368$, where
 $\exp(u) = e^u$. $P(X \leq 300) = \exp[-\exp(-1.6667)] = .828$,
 and $P(150 \leq X \leq 300) = .828 - .368 = .460$.
- b. The desired value c is the 90th percentile, so c satisfies
 $.9 = \exp\left[-\exp\left(\frac{-(c-150)}{90}\right)\right]$. Taking the natural log of each side twice in succession
 yields $\ln[\ln(.9)] = \frac{-(c-150)}{90}$, so $c = 90(2.250367) + 150 = 352.53$.
- c. $f(x) = F'(X) = \frac{1}{\beta} \cdot \exp\left[-\exp\left(\frac{-(x-\alpha)}{\beta}\right)\right] \cdot \exp\left(\frac{-(x-\alpha)}{\beta}\right)$
- d. We wish the value of x for which $f(x)$ is a maximum; this is the same as the value of x for which $\ln[f(x)]$ is a maximum. The equation of $\frac{d[\ln(f(x))]}{dx} = 0$ gives
 $\exp\left(\frac{-(x-\alpha)}{\beta}\right) = 1$, so $\frac{-(x-\alpha)}{\beta} = 0$, which implies that $x = \alpha$. Thus the mode is α .
- e. $E(X) = .5772\beta + \alpha = 201.95$, whereas the mode is 150 and the median is
 $-(90)\ln[-\ln(.5)] + 150 = 182.99$. The distribution is positively skewed.

104.

- a. $E(cX) = cE(X) = \frac{c}{\lambda}$
- b. $E[c(1 - .5e^{ax})] = \int_0^\infty c(1 - .5e^{ax}) \cdot \lambda e^{-\lambda x} dx = \frac{c[.5\lambda - a]}{\lambda - a}$

Chapter 4: Continuous Random Variables and Probability Distributions

105.

- a. From a graph of $f(x; \mu, \sigma)$ or by differentiation, $x^* = \mu$.
- b. No; the density function has constant height for $A \leq X \leq B$.
- c. $F(x; \lambda)$ is largest for $x = 0$ (the derivative at 0 does not exist since f is not continuous there) so $x^* = 0$.

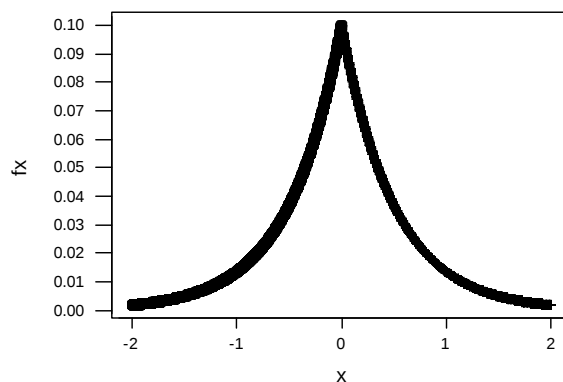
d. $\ln[f(x; \alpha, \beta)] = -\ln(\beta^\alpha) - \ln(\Gamma(\alpha)) + (\alpha - 1)\ln(x) - \frac{x}{\beta};$

$$\frac{d}{dx} \ln[f(x; \alpha, \beta)] = \frac{\alpha - 1}{x} - \frac{1}{\beta} \Rightarrow x = x^* = (\alpha - 1)\beta$$

e. From d $x^* = \left(\frac{\nu}{2} - 1\right)(2) = \nu - 2.$

106.

a. $\int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^0 .1e^{-2x}dx + \int_0^{\infty} .1e^{-2x}dx = .5 + .5 = 1$



b. For $x < 0$, $F(x) = \int_{-\infty}^x .1e^{-2y}dy = \frac{1}{2}e^{-2x}.$

For $x \geq 0$, $F(x) = \frac{1}{2} + \int_0^x .1e^{-2y}dy = 1 - \frac{1}{2}e^{-2x}.$

c. $P(X < 0) = F(0) = \frac{1}{2} = .5$, $P(X < 2) = F(2) = 1 - .5e^{-4} = .665$,

$P(-1 \leq X \leq 2) = F(2) - F(-1) = .256$, $1 - P(-2 \leq X \leq 2) = .670$

107.

a. Clearly $f(x; \lambda_1, \lambda_2, p) \geq 0$ for all x , and $\int_{-\infty}^{\infty} f(x; \lambda_1, \lambda_2, p) dx$

$$= \int_0^{\infty} [p\lambda_1 e^{-\lambda_1 x} + (1-p)\lambda_2 e^{-\lambda_2 x}] dx = p \int_0^{\infty} \lambda_1 e^{-\lambda_1 x} dx + (1-p) \int_0^{\infty} \lambda_2 e^{-\lambda_2 x} dx$$

$$= p + (1-p) = 1$$

b. For $x > 0$, $F(x; \lambda_1, \lambda_2, p) = \int_0^x f(y; \lambda_1, \lambda_2, p) dy = p(1 - e^{-\lambda_1 x}) + (1-p)(1 - e^{-\lambda_2 x})$.

c. $E(X) = \int_0^{\infty} x \cdot [p\lambda_1 e^{-\lambda_1 x} + (1-p)\lambda_2 e^{-\lambda_2 x}] dx$

$$= p \int_0^{\infty} x\lambda_1 e^{-\lambda_1 x} dx + (1-p) \int_0^{\infty} x\lambda_2 e^{-\lambda_2 x} dx = \frac{p}{\lambda_1} + \frac{(1-p)}{\lambda_2}$$

d. $E(X^2) = \frac{2p}{\lambda_1^2} + \frac{2(1-p)}{\lambda_2^2}$, so $\text{Var}(X) = \frac{2p}{\lambda_1^2} + \frac{2(1-p)}{\lambda_2^2} - \left[\frac{p}{\lambda_1} + \frac{(1-p)}{\lambda_2} \right]^2$

e. For an exponential r.v., $CV = \frac{1/\lambda}{1/\lambda} = 1$. For X hyperexponential,

$$CV = \left[\frac{\frac{2p}{\lambda_1^2} + \frac{2(1-p)}{\lambda_2^2}}{\left[\frac{p}{\lambda_1} + \frac{(1-p)}{\lambda_2} \right]^2} - 1 \right]^{1/2} = \left[\frac{2(p\lambda_2^2 + (1-p)\lambda_1^2)}{(p\lambda_2 + (1-p)\lambda_1)^2} - 1 \right]^{1/2}$$

$$= [2r - 1]^{1/2} \text{ where } r = \frac{(p\lambda_2^2 + (1-p)\lambda_1^2)}{(p\lambda_2 + (1-p)\lambda_1)^2}. \text{ But straightforward algebra shows that } r >$$

1 provided $\lambda_1 \neq \lambda_2$, so that $CV > 1$.

f. $\mu = \frac{n}{\lambda}$, $\sigma^2 = \frac{n}{\lambda^2}$, so $\sigma = \frac{\sqrt{n}}{\lambda}$ and $CV = \frac{1}{\sqrt{n}} < 1$ if $n > 1$.

108.

a. $1 = \int_5^{\infty} \frac{k}{x^{\alpha}} dx = k \cdot \frac{5^{1-\alpha}}{\alpha-1} \Rightarrow k = (\alpha-1)5^{1-\alpha}$ where we must have $\alpha > 1$.

b. For $x \geq 5$, $F(x) = \int_5^x \frac{k}{y^{\alpha}} dy = 5^{1-\alpha} \left[\frac{1}{5^{1-\alpha}} - \frac{1}{x^{\alpha-1}} \right] = 1 - \left(\frac{5}{x} \right)^{\alpha-1}$.

c. $E(X) = \int_5^{\infty} x \cdot \frac{k}{x^{\alpha}} dx = \int_5^{\infty} x \cdot \frac{k}{x^{\alpha-1}} dx = \frac{k}{5^{\alpha-2} \cdot (\alpha-2)}$, provided $\alpha > 2$.

d. $P\left(\ln\left(\frac{X}{5}\right) \leq y\right) = P\left(\frac{X}{5} \leq e^y\right) = P(X \leq 5e^y) = F(5e^y) = 1 - \left(\frac{5}{5e^y}\right)^{\alpha-1}$
 $1 - e^{-(\alpha-1)y}$, the cdf of an exponential r.v. with parameter $\alpha - 1$.

109.

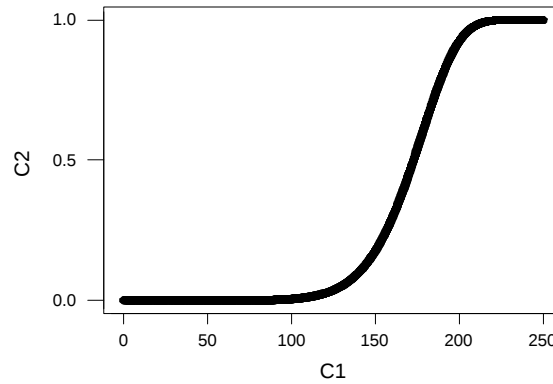
a. A lognormal distribution, since $\ln\left(\frac{I_o}{I_i}\right)$ is a normal r.v.

b. $P(I_o > 2I_i) = P\left(\frac{I_o}{I_i} > 2\right) = P\left(\ln\left(\frac{I_o}{I_i}\right) > \ln 2\right) = 1 - P\left(\ln\left(\frac{I_o}{I_i}\right) \leq \ln 2\right)$
 $1 - \Phi\left(\frac{\ln 2 - 1}{.05}\right) = 1 - \Phi(-6.14) = 1$

c. $E\left(\frac{I_o}{I_i}\right) = e^{1+.0025/2} = 2.72$, $Var\left(\frac{I_o}{I_i}\right) = e^{2+.0025} \cdot (e^{.0025} - 1) = .0185$

110.

a.



$$\begin{aligned} \text{b. } P(X > 175) &= 1 - F(175; 9, 180) = e^{-\left(\frac{175}{180}\right)^9} = .4602 \\ P(150 \leq X \leq 175) &= F(175; 9, 180) - F(150; 9, 180) \\ &= .5398 - .1762 = .3636 \end{aligned}$$

$$\text{c. } P(\text{at least one}) = 1 - P(\text{none}) = 1 - (1 - .3636)^2 = .5950$$

$$\text{d. } \text{We want the 10}^{\text{th}} \text{ percentile: } .10 = F(x; 9, 180) = 1 - e^{-\left(\frac{x}{180}\right)^9}. \text{ A small bit of algebra leads us to } x = 140.178. \text{ Thus 10\% of all tensile strengths will be less than 140.178 MPa.}$$

$$\begin{aligned} 111. \quad F(y) &= P(Y \leq y) = P(\sigma Z + \mu \leq y) = P\left(Z \leq \frac{(y - \mu)}{\sigma}\right) = \int_{-\infty}^{\frac{(y - \mu)}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz. \text{ Now} \\ &\text{differentiate with respect to } y \text{ to obtain a normal pdf with parameters } \mu \text{ and } \sigma. \end{aligned}$$

112.

$$\begin{aligned} \text{a. } F_Y(y) &= P(Y \leq y) = P(60X \leq y) = P\left(X \leq \frac{y}{60}\right) = F\left(\frac{y}{60\beta}; \alpha\right). \text{ Thus } f_Y(y) \\ &= f\left(\frac{y}{60\beta}; \alpha\right) \cdot \frac{1}{60\beta} = \frac{y^{\alpha-1} e^{-\frac{y}{60\beta}}}{(60\beta)^\alpha \Gamma(\alpha)}, \text{ which shows that } Y \text{ has a gamma distribution} \\ &\text{with parameters } \alpha \text{ and } 60\beta. \end{aligned}$$

$$\text{b. } \text{With } c \text{ replacing } 60 \text{ in } \text{a}, \text{ the same argument shows that } cX \text{ has a gamma distribution with parameters } \alpha \text{ and } c\beta.$$

Chapter 4: Continuous Random Variables and Probability Distributions

113.

- a. $Y = -\ln(X) \Rightarrow x = e^{-y} = k(y)$, so $k'(y) = -e^{-y}$. Thus since $f(x) = 1$, $g(y) = 1 \cdot |-e^{-y}| = e^{-y}$ for $0 < y < \infty$, so y has an exponential distribution with parameter $\lambda = 1$.
- b. $y = \sigma Z + \mu \Rightarrow y = h(z) = \sigma Z + \mu \Rightarrow z = k(y) = \frac{(y - \mu)}{\sigma}$ and $k'(y) = \frac{1}{\sigma}$, from which the result follows easily.
- c. $y = h(x) = cx \Rightarrow x = k(y) = \frac{y}{c}$ and $k'(y) = \frac{1}{c}$, from which the result follows easily.

114.

- a. If we let $\alpha = 2$ and $\beta = \sqrt{2}\sigma$, then we can manipulate $f(v)$ as follows:
- $$f(v) = \frac{v}{\sigma^2} e^{-v^2/2\sigma^2} = \frac{2}{2\sigma^2} v e^{-v^2/2\sigma^2} = \frac{2}{(\sqrt{2}\sigma)^2} v^{2-1} e^{-(v/\sqrt{2}\sigma)^2} = \frac{\alpha}{\beta^\alpha} v^{\alpha-1} e^{-(v/\beta)^\alpha},$$
- which is in the Weibull family of distributions.
- b. $F(v) = \int_0^v \frac{v}{400} e^{-\frac{v^2}{800}} dv$; cdf: $F(v; 2, \sqrt{2}\sigma) = 1 - e^{-\left(\frac{v}{\sqrt{2}\sigma}\right)^2} = 1 - e^{-\frac{v^2}{800}}$, so
- $$F(25; 2, \sqrt{2}) = 1 - e^{-\frac{625}{800}} = 1 - .458 = .542$$

115.

- a. Assuming independence, $P(\text{all 3 births occur on March 11}) = \left(\frac{1}{365}\right)^3 = .00000002$
- b. $\left(\frac{1}{365}\right)^3 (365) = .0000073$
- c. Let X = deviation from due date. $X \sim N(0, 19.88)$. Then the baby due on March 15 was 4 days early. $P(x = -4) \approx P(-4.5 < x < -3.5)$
- $$= \Phi\left(\frac{-3.5}{19.88}\right) - \Phi\left(\frac{-4.5}{19.88}\right) = \Phi(-.18) - \Phi(-.237) = .4286 - .4090 = .0196.$$
- Similarly, the baby due on April 1 was 21 days early, and $P(x = -21)$
- $$\approx \Phi\left(\frac{-20.5}{19.88}\right) - \Phi\left(\frac{-21.5}{19.88}\right) = \Phi(-1.03) - \Phi(-1.08) = .1515 - .1401 = .0114.$$
- The baby due on April 4 was 24 days early, and $P(x = -24) \approx .0097$
- Again, assuming independence, $P(\text{all 3 births occurred on March 11}) = (.0196)(.0114)(.0097) = .00002145$
- d. To calculate the probability of the three births happening on any day, we could make similar calculations as in part c for each possible day, and then add the probabilities.

Chapter 4: Continuous Random Variables and Probability Distributions

116.

- a. $F(x) = \lambda e^{-\lambda x}$ and $F(x) = 1 - e^{-\lambda x}$, so $r(x) = \frac{\lambda e^{-\lambda x}}{e^{-\lambda x}} = \lambda$, a constant (independent of X); this is consistent with the memoryless property of the exponential distribution.

- b. $r(x) = \left(\frac{\alpha}{\beta^\alpha}\right)x^{\alpha-1}$; for $\alpha > 1$ this is increasing, while for $\alpha < 1$ it is a decreasing function.

- c. $\ln(1 - F(x)) = -\int \alpha \left(1 - \frac{x}{\beta}\right) dx = -\alpha \left[x - \frac{x^2}{2\beta}\right] \Rightarrow F(x) = 1 - e^{-\alpha \left(x - \frac{x^2}{2\beta}\right)}$,

$$f(x) = \alpha \left(1 - \frac{x}{\beta}\right) e^{-\alpha \left(x - \frac{x^2}{2\beta}\right)} \quad 0 \leq x \leq \beta$$

117.

- a. $F_X(x) = P\left(-\frac{1}{\lambda} \ln(1 - U) \leq x\right) = P(\ln(1 - U) \geq -\lambda x) = P(1 - U \geq e^{-\lambda x})$
 $= P(U \leq 1 - e^{-\lambda x}) = 1 - e^{-\lambda x}$ since $F_U(u) = u$ (U is uniform on $[0, 1]$). Thus X has an exponential distribution with parameter λ .
- b. By taking successive random numbers $u_1, u_2, u_3, \dots$ and computing $x_i = -\frac{1}{10} \ln(1 - u_i)$,
 $\dots$ we obtain a sequence of values generated from an exponential distribution with parameter $\lambda = 10$.

118.

- a. $E(g(X)) \approx E[g(\mu) + g'(\mu)(X - \mu)] = E(g(\mu)) + g'(\mu) \cdot E(X - \mu)$, but $E(X - \mu) = 0$ and $E(g(\mu)) = g(\mu)$ (since $g(\mu)$ is constant), giving $E(g(X)) \approx g(\mu)$.
 $V(g(X)) \approx V[g(\mu) + g'(\mu)(X - \mu)] = V[g'(\mu)(X - \mu)] = (g'(\mu))^2 \cdot V(X - \mu) = (g'(\mu))^2 \cdot V(X)$.
- b. $g(I) = \frac{v}{I}, g'(I) = \frac{-v}{I^2}$, so $E(g(I)) = \mu_R \approx \frac{v}{\mu_I} = \frac{v}{20}$

$$V(g(I)) \approx \left(\frac{-v}{\mu_I^2}\right)^2 \cdot V(I), \sigma_{g(I)} \approx \frac{v}{20^2} \cdot \sigma_I = \frac{v}{800}$$

119. $g(\mu) + g'(\mu)(X - \mu) \leq g(X)$ implies that $E[g(\mu) + g'(\mu)(X - \mu)] = E(g(\mu)) = g(\mu) \leq E(g(X))$, i.e. that $g(E(X)) \leq E(g(X))$.

120. For $y > 0$, $F(y) = P(Y \leq y) = P\left(\frac{2X^2}{\beta^2} \leq y\right) = P\left(X^2 \leq \frac{\beta^2 y}{2}\right) = P\left(X \leq \frac{\beta\sqrt{y}}{\sqrt{2}}\right)$. Now

take the cdf of X (Weibull), replace x by $\frac{\beta\sqrt{y}}{\sqrt{2}}$, and then differentiate with respect to y to obtain the desired result $f_Y(y)$.

CHAPTER 5

Section 5.1

1.

- a. $P(X = 1, Y = 1) = p(1,1) = .20$
- b. $P(X \leq 1 \text{ and } Y \leq 1) = p(0,0) + p(0,1) + p(1,0) + p(1,1) = .42$
- c. At least one hose is in use at both islands. $P(X \neq 0 \text{ and } Y \neq 0) = p(1,1) + p(1,2) + p(2,1) + p(2,2) = .70$
- d. By summing row probabilities, $p_x(x) = .16, .34, .50$ for $x = 0, 1, 2$, and by summing column probabilities, $p_y(y) = .24, .38, .38$ for $y = 0, 1, 2$. $P(X \leq 1) = p_x(0) + p_x(1) = .50$
- e. $P(0,0) = .10$, but $p_x(0) \cdot p_y(0) = (.16)(.24) = .0384 \neq .10$, so X and Y are not independent.

2.

a.

		y				
p(x,y)		0	1	2	3	4
x	0	.30	.05	.025	.025	.10
	1	.18	.03	.015	.015	.06
	2	.12	.02	.01	.01	.04
		.6	.1	.05	.05	.2

- b. $P(X \leq 1 \text{ and } Y \leq 1) = p(0,0) + p(0,1) + p(1,0) + p(1,1) = .56$
 $= (.8)(.7) = P(X \leq 1) \cdot P(Y \leq 1)$
- c. $P(X + Y = 0) = P(X = 0 \text{ and } Y = 0) = p(0,0) = .30$
- d. $P(X + Y \leq 1) = p(0,0) + p(0,1) + p(1,0) = .53$

3.

- a. $p(1,1) = .15$, the entry in the 1st row and 1st column of the joint probability table.
- b. $P(X_1 = X_2) = p(0,0) + p(1,1) + p(2,2) + p(3,3) = .08 + .15 + .10 + .07 = .40$
- c. $A = \{ (x_1, x_2): x_1 \geq 2 + x_2 \} \cup \{ (x_1, x_2): x_2 \geq 2 + x_1 \}$
 $P(A) = p(2,0) + p(3,0) + p(4,0) + p(3,1) + p(4,1) + p(4,2) + p(0,2) + p(0,3) + p(1,3) = .22$
- d. $P(\text{exactly } 4) = p(1,3) + p(2,2) + p(3,1) + p(4,0) = .17$
 $P(\text{at least } 4) = P(\text{exactly } 4) + p(4,1) + p(4,2) + p(4,3) + p(3,2) + p(3,3) + p(2,3) = .46$

Chapter 5: Joint Probability Distributions and Random Samples

4.

- a. $P_1(0) = P(X_1 = 0) = p(0,0) + p(0,1) + p(0,2) + p(0,3) = .19$
 $P_1(1) = P(X_1 = 1) = p(1,0) + p(1,1) + p(1,2) + p(1,3) = .30$, etc.

x_1	0	1	2	3	4
$p_1(x_1)$	.19	.30	.25	.14	.12

- b. $P_2(0) = P(X_2 = 0) = p(0,0) + p(1,0) + p(2,0) + p(3,0) + p(4,0) = .19$, etc

x_2	0	1	2	3
$p_2(x_2)$	.19	.30	.28	.23

- c. $p(4,0) = 0$, yet $p_1(4) = .12 > 0$ and $p_2(0) = .19 > 0$, so $p(x_1, x_2) \neq p_1(x_1) \cdot p_2(x_2)$ for every (x_1, x_2) , and the two variables are not independent.

5.

- a. $P(X = 3, Y = 3) = P(3 \text{ customers, each with 1 package})$
 $= P(\text{each has 1 package} \mid 3 \text{ customers}) \cdot P(3 \text{ customers})$
 $= (.6)^3 \cdot (.25) = .054$
- b. $P(X = 4, Y = 11) = P(\text{total of 11 packages} \mid 4 \text{ customers}) \cdot P(4 \text{ customers})$
 Given that there are 4 customers, there are 4 different ways to have a total of 11 packages: 3, 3, 3, 2 or 3, 3, 2, 3 or 3, 2, 3, 3 or 2, 3, 3, 3. Each way has probability $(.1)^3(.3)$, so $p(4, 11) = 4(.1)^3(.3)(.15) = .00018$

6.

- a. $p(4,2) = P(Y = 2 \mid X = 4) \cdot P(X = 4) = \left[\binom{4}{2} (.6)^2 (.4)^2 \right] \cdot (.15) = .0518$
- b. $P(X = Y) = p(0,0) + p(1,1) + p(2,2) + p(3,3) + p(4,4) = .1 + (.2)(.6) + (.3)(.6)^2 + (.25)(.6)^3 + (.15)(.6)^4 = .4014$

Chapter 5: Joint Probability Distributions and Random Samples

- c. $p(x,y) = 0$ unless $y = 0, 1, \dots, x$; $x = 0, 1, 2, 3, 4$. For any such pair,

$$p(x,y) = P(Y = y | X = x) \cdot P(X = x) = \binom{x}{y} (.6)^y (.4)^{x-y} \cdot p_x(x)$$

$$p_y(4) = p(y = 4) = p(x = 4, y = 4) = p(4,4) = (.6)^4 (.15) = .0194$$

$$p_y(3) = p(3,3) + p(4,3) = (.6)^3 (.25) + \binom{4}{3} (.6)^3 (.4) (.15) = .1058$$

$$p_y(2) = p(2,2) + p(3,2) + p(4,2) = (.6)^2 (.3) + \binom{3}{2} (.6)^2 (.4) (.25) + \binom{4}{2} (.6)^2 (.4)^2 (.15) = .2678$$

$$p_y(1) = p(1,1) + p(2,1) + p(3,1) + p(4,1) = (.6)(.2) + \binom{2}{1} (.6)(.4)(.3) + \binom{3}{1} (.6)(.4)^2 (.25) + \binom{4}{1} (.6)(.4)^3 (.15) = .3590$$

$$p_y(0) = 1 - [.3590 + .2678 + .1058 + .0194] = .2480$$

7.

- a. $p(1,1) = .030$
- b. $P(X \leq 1 \text{ and } Y \leq 1) = p(0,0) + p(0,1) + p(1,0) + p(1,1) = .120$
- c. $P(X = 1) = p(1,0) + p(1,1) + p(1,2) = .100$; $P(Y = 1) = p(0,1) + \dots + p(5,1) = .300$
- d. $P(\text{overflow}) = P(X + 3Y > 5) = 1 - P(X + 3Y \leq 5) = 1 - P[(X,Y) = (0,0) \text{ or } \dots \text{ or } (5,0) \text{ or } (0,1) \text{ or } (1,1) \text{ or } (2,1)] = 1 - .620 = .380$
- e. The marginal probabilities for X (row sums from the joint probability table) are $p_x(0) = .05$, $p_x(1) = .10$, $p_x(2) = .25$, $p_x(3) = .30$, $p_x(4) = .20$, $p_x(5) = .10$; those for Y (column sums) are $p_y(0) = .5$, $p_y(1) = .3$, $p_y(2) = .2$. It is now easily verified that for every (x,y) , $p(x,y) = p_x(x) \cdot p_y(y)$, so X and Y are independent.

8.

$$\text{a. numerator} = \binom{8}{3} \binom{10}{2} \binom{12}{1} = (56)(45)(12) = 30,240$$

$$\text{denominator} = \binom{30}{6} = 593,775; p(3,2) = \frac{30,240}{593,775} = .0509$$

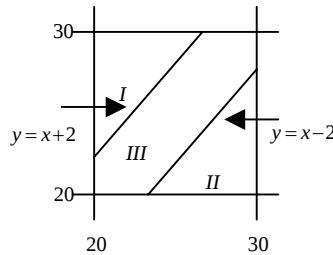
$$\text{b. } p(x,y) = \begin{cases} \frac{\binom{8}{x} \binom{10}{y} \binom{12}{6-(x+y)}}{\binom{30}{6}} & \begin{array}{l} x, y \text{ are non-negative} \\ \text{integers such that} \\ 0 \leq x+y \leq 6 \end{array} \\ 0 & \text{otherwise} \end{cases}$$

9.

$$\begin{aligned} \text{a. } 1 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y) dx dy = \int_{20}^{30} \int_{20}^{30} K(x^2 + y^2) dx dy \\ &= K \int_{20}^{30} \int_{20}^{30} x^2 dy dx + K \int_{20}^{30} \int_{20}^{30} y^2 dx dy = 10K \int_{20}^{30} x^2 dx + 10K \int_{20}^{30} y^2 dy \\ &= 20K \cdot \left(\frac{19,000}{3} \right) \Rightarrow K = \frac{3}{380,000} \end{aligned}$$

$$\begin{aligned} \text{b. } P(X < 26 \text{ and } Y < 26) &= \int_{20}^{26} \int_{20}^{26} K(x^2 + y^2) dx dy = 12K \int_{20}^{26} x^2 dx \\ &= 4Kx^3 \Big|_{20}^{26} = 38,304K = .3024 \end{aligned}$$

c.



$$\begin{aligned} P(|X - Y| \leq 2) &= \iint_{\text{region III}} f(x,y) dx dy \\ &= 1 - \iint_I f(x,y) dx dy - \iint_{II} f(x,y) dx dy \\ &= 1 - \int_{20}^{28} \int_{x+2}^{30} f(x,y) dy dx - \int_{22}^{30} \int_{20}^{x-2} f(x,y) dy dx \\ &= (\text{after much algebra}) .3593 \end{aligned}$$

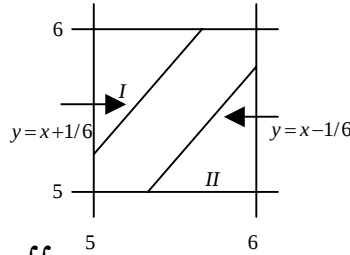
Chapter 5: Joint Probability Distributions and Random Samples

- d. $f_x(x) = \int_{-\infty}^{\infty} f(x, y) dy = \int_{20}^{30} K(x^2 + y^2) dy = 10Kx^2 + K \frac{y^3}{3} \Big|_{20}^{30}$
 $= 10Kx^2 + .05, \quad 20 \leq x \leq 30$
- e. $f_y(y)$ is obtained by substituting y for x in (d); clearly $f(x, y) \neq f_x(x) \cdot f_y(y)$, so X and Y are not independent.

10.

- a. $f(x, y) = \begin{cases} 1 & 5 \leq x \leq 6, 5 \leq y \leq 6 \\ 0 & \text{otherwise} \end{cases}$
 since $f_x(x) = 1, f_y(y) = 1$ for $5 \leq x \leq 6, 5 \leq y \leq 6$
- b. $P(5.25 \leq X \leq 5.75, 5.25 \leq Y \leq 5.75) = P(5.25 \leq X \leq 5.75) \cdot P(5.25 \leq Y \leq 5.75) =$ (by independence) $(.5)(.5) = .25$

c.



$$P((X, Y) \in A) = \iint_A 1 dx dy$$

$$= \text{area of } A = 1 - (\text{area of I} + \text{area of II})$$

$$= 1 - \frac{25}{36} = \frac{11}{36} = .306$$

11.

- a. $p(x, y) = \frac{e^{-\lambda} \lambda^x}{x!} \cdot \frac{e^{-\mu} \mu^y}{y!}$ for $x = 0, 1, 2, \dots; y = 0, 1, 2, \dots$
- b. $p(0, 0) + p(0, 1) + p(1, 0) = e^{-\lambda-\mu} [1 + \lambda + \mu]$
- c. $P(X+Y=m) = \sum_{k=0}^m P(X=k, Y=m-k) = \sum_{k=0}^m e^{-\lambda-\mu} \frac{\lambda^k}{k!} \frac{\mu^{m-k}}{(m-k)!}$

$$\frac{e^{-(\lambda+\mu)}}{m!} \sum_{k=0}^m \binom{m}{k} \lambda^k \mu^{m-k} = \frac{e^{-(\lambda+\mu)} (\lambda + \mu)^m}{m!},$$
 so the total # of errors $X+Y$ also has a Poisson distribution with parameter $\lambda + \mu$.

12.

a.
$$P(X > 3) = \int_3^\infty \int_0^\infty x e^{-x(1+y)} dy dx = \int_3^\infty e^{-x} dx = .050$$

b. The marginal pdf of X is $\int_0^\infty x e^{-x(1+y)} dy = e^{-x}$ for $0 \leq x$; that of Y is

$$\int_3^\infty x e^{-x(1+y)} dx = \frac{1}{(1+y)^2} \text{ for } 0 \leq y. \text{ It is now clear that } f(x,y) \text{ is not the product of the marginal pdf's, so the two r.v's are not independent.}$$

c.
$$\begin{aligned} P(\text{at least one exceeds } 3) &= 1 - P(X \leq 3 \text{ and } Y \leq 3) \\ &= 1 - \int_0^3 \int_0^3 x e^{-x(1+y)} dy dx = 1 - \int_0^3 \int_0^3 x e^{-x} e^{-xy} dy \\ &= 1 - \int_0^3 e^{-x} (1 - e^{-3x}) dx = e^{-3} + .25 - .25e^{-12} = .300 \end{aligned}$$

13.

a.
$$f(x,y) = f_x(x) \cdot f_y(y) = \begin{cases} e^{-x-y} & x \geq 0, y \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

b.
$$P(X \leq 1 \text{ and } Y \leq 1) = P(X \leq 1) \cdot P(Y \leq 1) = (1 - e^{-1})(1 - e^{-1}) = .400$$

c.
$$\begin{aligned} P(X + Y \leq 2) &= \int_0^2 \int_0^{2-x} e^{-x-y} dy dx = \int_0^2 e^{-x} [1 - e^{-(2-x)}] dx \\ &= \int_0^2 (e^{-x} - e^{-2}) dx = 1 - e^{-2} - 2e^{-2} = .594 \end{aligned}$$

d.
$$P(X + Y \leq 1) = \int_0^1 e^{-x} [1 - e^{-(1-x)}] dx = 1 - 2e^{-1} = .264,$$

so $P(1 \leq X + Y \leq 2) = P(X + Y \leq 2) - P(X + Y \leq 1) = .594 - .264 = .330$

14.

a.
$$P(X_1 < t, X_2 < t, \dots, X_{10} < t) = P(X_1 < t) \dots P(X_{10} < t) = (1 - e^{-\lambda t})^{10}$$

b. If "success" = {fail before t}, then $p = P(\text{success}) = 1 - e^{-\lambda t}$,
and $P(k \text{ successes among } 10 \text{ trials}) = \binom{10}{k} (1 - e^{-\lambda t})^k (e^{-\lambda t})^{10-k}$

c.
$$P(\text{exactly } 5 \text{ fail}) = P(5 \text{ of } \lambda \text{'s fail and other } 5 \text{ don't}) + P(4 \text{ of } \lambda \text{'s fail, } \mu \text{ fails, and other } 5 \text{ don't})$$

$$= \binom{9}{5} (1 - e^{-\lambda t})^5 (e^{-\lambda t})^4 (e^{-\mu t}) + \binom{9}{4} (1 - e^{-\lambda t})^4 (1 - e^{-\mu t}) (e^{-\lambda t})^5$$

15.

$$\begin{aligned} \text{a. } F(y) &= P(Y \leq y) = P[(X_1 \leq y) \cup ((X_2 \leq y) \cap (X_3 \leq y))] \\ &= P(X_1 \leq y) + P[(X_2 \leq y) \cap (X_3 \leq y)] - P[(X_1 \leq y) \cap (X_2 \leq y) \cap (X_3 \leq y)] \\ &= (1 - e^{-\lambda y}) + (1 - e^{-\lambda y})^2 - (1 - e^{-\lambda y})^3 \text{ for } y \geq 0 \end{aligned}$$

$$\begin{aligned} \text{b. } f(y) &= F'(y) = \lambda e^{-\lambda y} + 2(1 - e^{-\lambda y})(\lambda e^{-\lambda y}) - 3(1 - e^{-\lambda y})^2(\lambda e^{-\lambda y}) \\ &= 4\lambda e^{-2\lambda y} - 3\lambda e^{-3\lambda y} \text{ for } y \geq 0 \end{aligned}$$

$$E(Y) = \int_0^{\infty} y \cdot (4\lambda e^{-2\lambda y} - 3\lambda e^{-3\lambda y}) dy = 2\left(\frac{1}{2\lambda}\right) - \frac{1}{3\lambda} = \frac{2}{3\lambda}$$

16.

$$\begin{aligned} \text{a. } f(x_1, x_3) &= \int_{-\infty}^{\infty} f(x_1, x_2, x_3) dx_2 = \int_0^{1-x_1-x_3} kx_1x_2(1-x_3) dx_2 \\ &= 72x_1(1-x_3)(1-x_1-x_3)^2 \quad 0 \leq x_1, 0 \leq x_3, x_1 + x_3 \leq 1 \end{aligned}$$

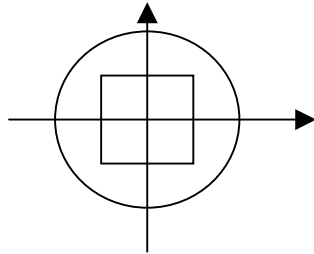
$$\begin{aligned} \text{b. } P(X_1 + X_3 \leq .5) &= \int_0^{.5} \int_0^{.5-x_1} 72x_1(1-x_3)(1-x_1-x_3)^2 dx_2 dx_1 \\ &= (\text{after much algebra}) .53125 \end{aligned}$$

$$\begin{aligned} \text{c. } f_{x_1}(x_1) &= \int_{-\infty}^{\infty} f(x_1, x_3) dx_3 = \int 72x_1(1-x_3)(1-x_1-x_3)^2 dx_3 \\ &= 18x_1 - 48x_1^2 + 36x_1^3 - 6x_1^5 \quad 0 \leq x_1 \leq 1 \end{aligned}$$

17.

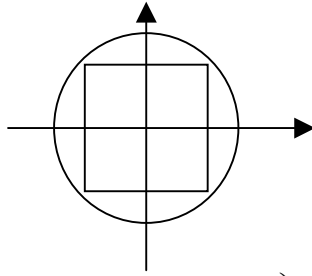
$$\begin{aligned} \text{a. } P((X, Y) \text{ within a circle of radius } \frac{R}{2}) &= P(A) = \iint_A f(x, y) dx dy \\ &= \frac{1}{\pi R^2} \iint_A dx dy = \frac{\text{area of } A}{\pi R^2} = \frac{1}{4} = .25 \end{aligned}$$

b.



$$P\left(-\frac{R}{2} \leq X \leq \frac{R}{2}, -\frac{R}{2} \leq Y \leq \frac{R}{2}\right) = \frac{R^2}{\pi R^2} = \frac{1}{\pi}$$

c.



$$P\left(-\frac{R}{\sqrt{2}} \leq X \leq \frac{R}{\sqrt{2}}, -\frac{R}{\sqrt{2}} \leq Y \leq \frac{R}{\sqrt{2}}\right) = \frac{2R^2}{\pi R^2} = \frac{2}{\pi}$$

- d. $f_x(x) = \int_{-\infty}^{\infty} f(x, y) dy = \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} \frac{1}{\pi R^2} dy = \frac{2\sqrt{R^2-x^2}}{\pi R^2}$ for $-R \leq x \leq R$ and similarly for $f_y(y)$. X and Y are not independent since e.g. $f_x(.9R) = f_y(.9R) > 0$, yet $f(.9R, .9R) = 0$ since $(.9R, .9R)$ is outside the circle of radius R .

18.

- a. $P_{y|x}(y|1)$ results from dividing each entry in $x = 1$ row of the joint probability table by $p_x(1) = .34$:

$$P_{y|x}(0|1) = \frac{.08}{.34} = .2353$$

$$P_{y|x}(1|1) = \frac{.20}{.34} = .5882$$

$$P_{y|x}(2|1) = \frac{.06}{.34} = .1765$$

- b. $P_{y|x}(x|2)$ is requested; to obtain this divide each entry in the $y = 2$ row by $p_x(2) = .50$:

y	0	1	2
$P_{y x}(y 2)$	.12	.28	.60

- c. $P(Y \leq 1 | x = 2) = P_{y|x}(0|2) + P_{y|x}(1|2) = .12 + .28 = .40$

- d. $P_{x|y}(x|2)$ results from dividing each entry in the $y = 2$ column by $p_y(2) = .38$:

x	0	1	2
$P_{x y}(x 2)$	.0526	.1579	.7895

19.

$$\begin{aligned} \text{a. } f_{Y|X}(y|x) &= \frac{f(x,y)}{f_X(x)} = \frac{k(x^2 + y^2)}{10kx^2 + .05} & 20 \leq y \leq 30 \\ f_{X|Y}(x|y) &= \frac{k(x^2 + y^2)}{10ky^2 + .05} & 20 \leq x \leq 30 \quad \left(k = \frac{3}{380,000} \right) \end{aligned}$$

$$\begin{aligned} \text{b. } P(Y \geq 25 | X = 22) &= \int_{25}^{30} f_{Y|X}(y|22) dy \\ &= \int_{25}^{30} \frac{k((22)^2 + y^2)}{10k(22)^2 + .05} dy = .783 \end{aligned}$$

$$P(Y \geq 25) = \int_{25}^{30} f_Y(y) dy = \int_{25}^{30} (10ky^2 + .05) dy = .75$$

$$\begin{aligned} \text{c. } E(Y | X=22) &= \int_{-\infty}^{\infty} y \cdot f_{Y|X}(y|22) dy = \int_{20}^{30} y \cdot \frac{k((22)^2 + y^2)}{10k(22)^2 + .05} dy \\ &= 25.372912 \\ E(Y^2 | X=22) &= \int_{20}^{30} y^2 \cdot \frac{k((22)^2 + y^2)}{10k(22)^2 + .05} dy = 652.028640 \\ V(Y | X = 22) &= E(Y^2 | X=22) - [E(Y | X=22)]^2 = 8.243976 \end{aligned}$$

20.

$$\begin{aligned} \text{a. } f_{x_3|x_1, x_2}(x_3 | x_1, x_2) &= \frac{f(x_1, x_2, x_3)}{f_{x_1, x_2}(x_1, x_2)} \text{ where } f_{x_1, x_2}(x_1, x_2) = \text{the marginal joint pdf} \\ \text{of } (X_1, X_2) &= \int_{-\infty}^{\infty} f(x_1, x_2, x_3) dx_3 \end{aligned}$$

$$\begin{aligned} \text{b. } f_{x_2, x_3|x_1}(x_2, x_3 | x_1) &= \frac{f(x_1, x_2, x_3)}{f_{x_1}(x_1)} \text{ where} \\ f_{x_1}(x_1) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1, x_2, x_3) dx_2 dx_3 \end{aligned}$$

21. For every x and y, $f_{Y|X}(y|x) = f_Y(y)$, since then $f(x,y) = f_{Y|X}(y|x) \cdot f_X(x) = f_Y(y) \cdot f_X(x)$, as required.

Section 5.2

22.

$$\begin{aligned} \text{a. } E(X + Y) &= \sum_x \sum_y (x + y) p(x, y) = (0 + 0)(.02) \\ &\quad + (0 + 5)(.06) + \dots + (10 + 15)(.01) = 14.10 \end{aligned}$$

$$\begin{aligned} \text{b. } E[\max(X, Y)] &= \sum_x \sum_y \max(x, y) \cdot p(x, y) \\ &= (0)(.02) + (5)(.06) + \dots + (15)(.01) = 9.60 \end{aligned}$$

$$\begin{aligned} 23. \quad E(X_1 - X_2) &= \sum_{x_1=0}^4 \sum_{x_2=0}^3 (x_1 - x_2) \cdot p(x_1, x_2) = \\ &= (0 - 0)(.08) + (0 - 1)(.07) + \dots + (4 - 3)(.06) = .15 \\ &\quad \text{(which also equals } E(X_1) - E(X_2) = 1.70 - 1.55) \end{aligned}$$

24. Let $h(X, Y)$ = # of individuals who handle the message.

		y					
h(x,y)		1	2	3	4	5	6
x	1	-	2	3	4	3	2
	2	2	-	2	3	4	3
	3	3	2	-	2	3	4
	4	4	3	2	-	2	3
	5	3	4	3	2	-	2
	6	2	3	4	3	2	-

$$\text{Since } p(x, y) = \frac{1}{30} \text{ for each possible } (x, y), E[h(X, Y)] = \sum_x \sum_y h(x, y) \cdot \frac{1}{30} = \frac{84}{30} = 2.80$$

$$25. \quad E(XY) = E(X) \cdot E(Y) = L \cdot L = L^2$$

$$\begin{aligned} 26. \quad \text{Revenue} &= 3X + 10Y, \text{ so } E(\text{revenue}) = E(3X + 10Y) \\ &= \sum_{x=0}^5 \sum_{y=0}^2 (3x + 10y) \cdot p(x, y) = 0 \cdot p(0, 0) + \dots + 35 \cdot p(5, 2) = 15.4 \end{aligned}$$

$$27. \quad E[h(X,Y)] = \int_0^1 \int_0^1 |x-y| \cdot 6x^2 y dx dy = 2 \int_0^1 \int_0^x (x-y) \cdot 6x^2 y dy dx$$

$$12 \int_0^1 \int_0^x (x^3 y - x^2 y^2) dy dx = 12 \int_0^1 \frac{x^5}{6} dx = \frac{1}{3}$$

$$28. \quad E(XY) = \sum_x \sum_y xy \cdot p(x, y) = \sum_x \sum_y xy \cdot p_x(x) \cdot p_y(y) = \sum_x xp_x(x) \cdot \sum_y yp_y(y) \\ = E(X) \cdot E(Y). \quad (\text{replace } \Sigma \text{ with } \int \text{ in the continuous case})$$

$$29. \quad \text{Cov}(X,Y) = -\frac{2}{75} \text{ and } \mu_x = \mu_y = \frac{2}{5}. \quad E(X^2) = \int_0^1 x^2 \cdot f_x(x) dx \\ = 12 \int_0^1 x^3 (1-x^2) dx = \frac{12}{60} = \frac{1}{5}, \text{ so } \text{Var}(X) = \frac{1}{5} - \frac{4}{25} = \frac{1}{25} \\ \text{Similarly, } \text{Var}(Y) = \frac{1}{25}, \text{ so } \rho_{X,Y} = \frac{-\frac{2}{75}}{\sqrt{\frac{1}{25}} \cdot \sqrt{\frac{1}{25}}} = -\frac{50}{75} = -.667$$

30.

$$\text{a. } E(X) = 5.55, E(Y) = 8.55, E(XY) = (0)(.02) + (0)(.06) + \dots + (150)(.01) = 44.25, \text{ so} \\ \text{Cov}(X,Y) = 44.25 - (5.55)(8.55) = -3.20$$

$$\text{b. } \sigma_X^2 = 12.45, \sigma_Y^2 = 19.15, \text{ so } \rho_{X,Y} = \frac{-3.20}{\sqrt{(12.45)(19.15)}} = -.207$$

31.

$$\text{a. } E(X) = \int_{20}^{30} x f_x(x) dx = \int_{20}^{30} x [10Kx^2 + .05] dx = 25.329 = E(Y)$$

$$E(XY) = \int_{20}^{30} \int_{20}^{30} xy \cdot K(x^2 + y^2) dx dy = 641.447$$

$$\Rightarrow \text{Cov}(X,Y) = 641.447 - (25.329)^2 = -.111$$

$$\text{b. } E(X^2) = \int_{20}^{30} x^2 [10Kx^2 + .05] dx = 649.8246 = E(Y^2),$$

$$\text{so } \text{Var}(X) = \text{Var}(Y) = 649.8246 - (25.329)^2 = 8.2664$$

$$\Rightarrow \rho = \frac{-.111}{\sqrt{(8.2664)(8.2664)}} = -.0134$$

Chapter 5: Joint Probability Distributions and Random Samples

- 32.** There is a difficulty here. Existence of ρ requires that both X and Y have finite means and variances. Yet since the marginal pdf of Y is $\frac{1}{(1-y)^2}$ for $y \geq 0$,
- $$E(y) = \int_0^\infty \frac{y}{(1+y)^2} dy = \int_0^\infty \frac{(1+y-1)}{(1+y)^2} dy = \int_0^\infty \frac{1}{(1+y)} dy - \int_0^\infty \frac{1}{(1+y)^2} dy,$$
- and the first integral is not finite. Thus ρ itself is undefined.
- 33.** Since $E(XY) = E(X) \cdot E(Y)$, $\text{Cov}(X, Y) = E(XY) - E(X) \cdot E(Y) = E(X) \cdot E(Y) - E(X) \cdot E(Y) = 0$, and since $\text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_x \sigma_y}$, then $\text{Corr}(X, Y) = 0$
- 34.**
- a.** In the discrete case, $\text{Var}[h(X, Y)] = E\{[h(X, Y) - E(h(X, Y))]^2\} = \sum_x \sum_y [h(x, y) - E(h(X, Y))]^2 p(x, y) = \sum_x \sum_y [h(x, y)^2 p(x, y)] - [E(h(X, Y))]^2$
with $\iint$ replacing $\sum \sum$ in the continuous case.
- b.** $E[h(X, Y)] = E[\max(X, Y)] = 9.60$, and $E[h^2(X, Y)] = E[(\max(X, Y))^2] = (0)^2(.02) + (5)^2(.06) + \dots + (15)^2(.01) = 105.5$, so $\text{Var}[\max(X, Y)] = 105.5 - (9.60)^2 = 13.34$
- 35.**
- a.** $\text{Cov}(aX + b, cY + d) = E[(aX + b)(cY + d)] - E(aX + b) \cdot E(cY + d)$
 $= E[acXY + adX + bcY + bd] - (aE(X) + b)(cE(Y) + d)$
 $= acE(XY) - acE(X)E(Y) = ac\text{Cov}(X, Y)$
- b.** $\text{Corr}(aX + b, cY + d) = \frac{\text{Cov}(aX + b, cY + d)}{\sqrt{\text{Var}(aX + b)} \sqrt{\text{Var}(cY + d)}} = \frac{ac\text{Cov}(X, Y)}{|a| \cdot |c| \sqrt{\text{Var}(X) \cdot \text{Var}(Y)}}$
 $= \text{Corr}(X, Y)$ when a and c have the same signs.
- c.** When a and c differ in sign, $\text{Corr}(aX + b, cY + d) = -\text{Corr}(X, Y)$.
- 36.** $\text{Cov}(X, Y) = \text{Cov}(X, aX + b) = E[X \cdot (aX + b)] - E(X) \cdot E(aX + b) = a \text{Var}(X)$,
 so $\text{Corr}(X, Y) = \frac{a\text{Var}(X)}{\sqrt{\text{Var}(X) \cdot \text{Var}(Y)}} = \frac{a\text{Var}(X)}{\sqrt{\text{Var}(X) \cdot a^2\text{Var}(X)}} = 1$ if $a > 0$, and -1 if $a < 0$

Section 5.3

37.

	$P(x_1)$	.20	.50	.30
$P(x_2)$	$x_2 \mid x_1$	25	40	65
.20	25	.04	.10	.06
.50	40	.10	.25	.15
.30	65	.06	.15	.09

a.

$\bar{x}$	25	32.5	40	45	52.5	65
$p(\bar{x})$	.04	.20	.25	.12	.30	.09

$$E(\bar{x}) = (25)(.04) + 32.5(.20) + \dots + 65(.09) = 44.5 = \mu$$

b.

s^2	0	112.5	312.5	800
$P(s^2)$	.38	.20	.30	.12

$$E(s^2) = 212.25 = \sigma^2$$

38.

a.

T_0	0	1	2	3	4
$P(T_0)$	.04	.20	.37	.30	.09

$$\text{b. } \mu_{T_0} = E(T_0) = 2.2 = 2 \cdot \mu$$

$$\text{c. } \sigma_{T_0}^2 = E(T_0^2) - E(T_0)^2 = 5.82 - (2.2)^2 = .98 = 2 \cdot \sigma^2$$

Chapter 5: Joint Probability Distributions and Random Samples

39.

x	0	1	2	3	4	5	6	7	8	9	10
x/n	0	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0
p(x/n)	.000	.000	.000	.001	.005	.027	.088	.201	.302	.269	.107

X is a binomial random variable with $p = .8$.

40.

- a. Possible values of M are: 0, 5, 10. $M = 0$ when all 3 envelopes contain 0 money, hence $p(M = 0) = (.5)^3 = .125$. $M = 10$ when there is a single envelope with \$10, hence $p(M = 10) = 1 - p(\text{no envelopes with } \$10) = 1 - (.8)^3 = .488$.
 $p(M = 5) = 1 - [.125 + .488] = .387$.

M	0	5	10
p(M)	.125	.387	.488

An alternative solution would be to list all 27 possible combinations using a tree diagram and computing probabilities directly from the tree.

- b. The statistic of interest is M, the maximum of x_1 , x_2 , or x_3 , so that $M = 0, 5$, or 10 . The population distribution is as follows:

x	0	5	10
p(x)	1/2	3/10	1/5

Write a computer program to generate the digits 0 – 9 from a uniform distribution. Assign a value of 0 to the digits 0 – 4, a value of 5 to digits 5 – 7, and a value of 10 to digits 8 and 9. Generate samples of increasing sizes, keeping the number of replications constant and compute M from each sample. As n, the sample size, increases, $p(M = 0)$ goes to zero, $p(M = 10)$ goes to one. Furthermore, $p(M = 5)$ goes to zero, but at a slower rate than $p(M = 0)$.

Chapter 5: Joint Probability Distributions and Random Samples

41.

Outcome	1,1	1,2	1,3	1,4	2,1	2,2	2,3	2,4
Probability	.16	.12	.08	.04	.12	.09	.06	.03
$\bar{x}$	1	1.5	2	2.5	1.5	2	2.5	3
r	0	1	2	3	1	0	1	2

Outcome	3,1	3,2	3,3	3,4	4,1	4,2	4,3	4,4
Probability	.08	.06	.04	.02	.04	.03	.02	.01
$\bar{x}$	2	2.5	3	3.5	2.5	3	3.5	4
r	2	1	0	1	3	2	1	2

a.

$\bar{x}$	1	1.5	2	2.5	3	3.5	4
$p(\bar{x})$	.16	.24	.25	.20	.10	.04	.01

b. $P(\bar{x} \leq 2.5) = .8$

c.

r	0	1	2	3
p(r)	.30	.40	.22	.08

d. $P(\bar{X} \leq 1.5) = P(1,1,1,1) + P(2,1,1,1) + \dots + P(1,1,1,2) + P(1,1,2,2) + \dots + P(2,2,1,1) + P(3,1,1,1) + \dots + P(1,1,1,3)$
 $= (.4)^4 + 4(.4)^3(.3) + 6(.4)^2(.3)^2 + 4(.4)^2(.2)^2 = .2400$

42.

a.

$\bar{x}$	27.75	28.0	29.7	29.95	31.65	31.9	33.6
$p(\bar{x})$	$\frac{4}{30}$	$\frac{2}{30}$	$\frac{6}{30}$	$\frac{4}{30}$	$\frac{8}{30}$	$\frac{4}{30}$	$\frac{2}{30}$

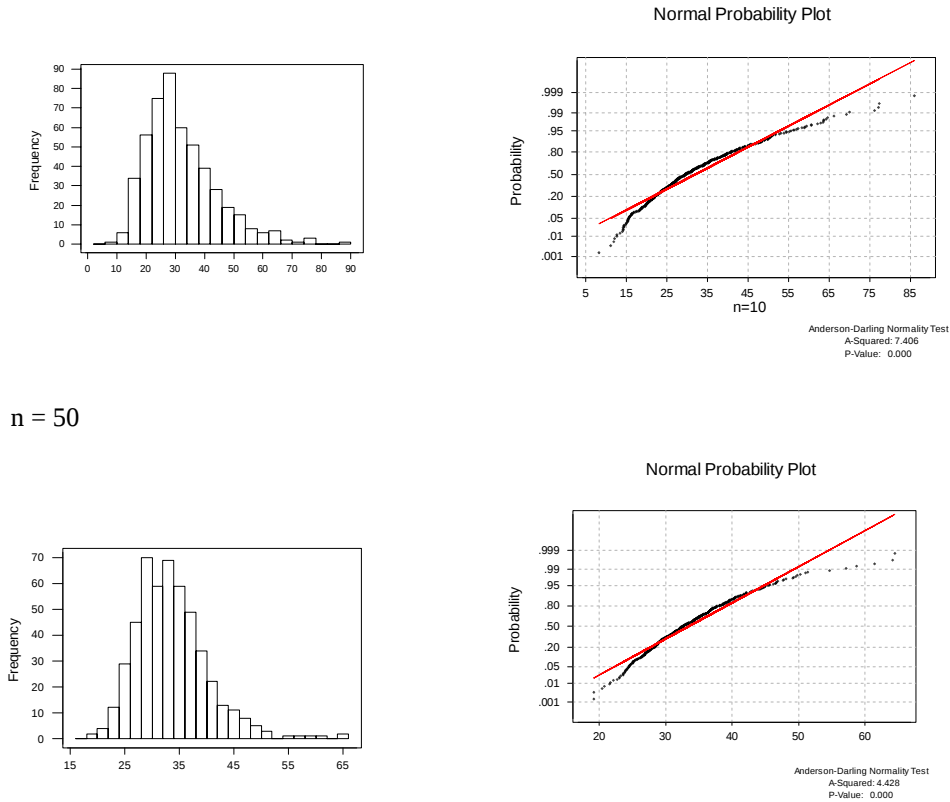
b.

$\bar{x}$	27.75	31.65	31.9
$p(\bar{x})$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

c. all three values are the same: 30.4333

Chapter 5: Joint Probability Distributions and Random Samples

43. The statistic of interest is the fourth spread, or the difference between the medians of the upper and lower halves of the data. The population distribution is uniform with $A = 8$ and $B = 10$. Use a computer to generate samples of sizes $n = 5, 10, 20$, and 30 from a uniform distribution with $A = 8$ and $B = 10$. Keep the number of replications the same (say 500, for example). For each sample, compute the upper and lower fourth, then compute the difference. Plot the sampling distributions on separate histograms for $n = 5, 10, 20$, and 30 .
44. Use a computer to generate samples of sizes $n = 5, 10, 20$, and 30 from a Weibull distribution with parameters as given, keeping the number of replications the same, as in problem 43 above. For each sample, calculate the mean. Below is a histogram, and a normal probability plot for the sampling distribution of $\bar{X}$ for $n = 5$, both generated by Minitab. This sampling distribution appears to be normal, so since larger sample sizes will produce distributions that are closer to normal, the others will also appear normal.
45. Using Minitab to generate the necessary sampling distribution, we can see that as n increases, the distribution slowly moves toward normality. However, even the sampling distribution for $n = 50$ is not yet approximately normal.



Section 5.4

46. $\mu = 12 \text{ cm}$ $\sigma = .04 \text{ cm}$

a. $n = 16$ $E(\bar{X}) = \mu = 12 \text{ cm}$ $\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}} = \frac{.04}{4} = .01 \text{ cm}$

b. $n = 64$ $E(\bar{X}) = \mu = 12 \text{ cm}$ $\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}} = \frac{.04}{8} = .005 \text{ cm}$

c. $\bar{X}$ is more likely to be within .01 cm of the mean (12 cm) with the second, larger, sample. This is due to the decreased variability of $\bar{X}$ with a larger sample size.

47. $\mu = 12 \text{ cm}$ $\sigma = .04 \text{ cm}$

a. $n = 16$ $P(11.99 \leq \bar{X} \leq 12.01) = P\left(\frac{11.99 - 12}{.01} \leq Z \leq \frac{12.01 - 12}{.01}\right)$
 $= P(-1 \leq Z \leq 1)$
 $= \Phi(1) - \Phi(-1)$
 $= .8413 - .1587$
 $= .6826$

b. $n = 25$ $P(\bar{X} > 12.01) = P\left(Z > \frac{12.01 - 12}{.04/5}\right) = P(Z > 1.25)$
 $= 1 - \Phi(1.25)$
 $= 1 - .8944$
 $= .1056$

48.

a. $\mu_{\bar{X}} = \mu = 50$, $\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}} = \frac{1}{\sqrt{100}} = .10$

$P(49.75 \leq \bar{X} \leq 50.25) = P\left(\frac{49.75 - 50}{.10} \leq Z \leq \frac{50.25 - 50}{.10}\right)$
 $= P(-2.5 \leq Z \leq 2.5) = .9876$

b. $P(49.75 \leq \bar{X} \leq 50.25) \approx P\left(\frac{49.75 - 49.8}{.10} \leq Z \leq \frac{50.25 - 49.8}{.10}\right)$
 $= P(-.5 \leq Z \leq 4.5) = .6915$

Chapter 5: Joint Probability Distributions and Random Samples

49.

- a. 11 P.M. – 6:50 P.M. = 250 minutes. With $T_0 = X_1 + \dots + X_{40}$ = total grading time,
 $\mu_{T_0} = n\mu = (40)(6) = 240$ and $\sigma_{T_0} = \sigma\sqrt{n} = 37.95$, so $P(T_0 \leq 250) \approx$

$$P\left(Z \leq \frac{250 - 240}{37.95}\right) = P(Z \leq .26) = .6026$$

- b. $P(T_0 > 260) = P\left(Z > \frac{260 - 240}{37.95}\right) = P(Z > .53) = .2981$

50.

$\mu = 10,000$ psi $\sigma = 500$ psi

- a. $n = 40$

$$\begin{aligned} P(9,900 \leq \bar{X} \leq 10,200) &\approx P\left(\frac{9,900 - 10,000}{500/\sqrt{40}} \leq Z \leq \frac{10,200 - 10,000}{500/\sqrt{40}}\right) \\ &= P(-1.26 \leq Z \leq 2.53) \\ &= \Phi(2.53) - \Phi(-1.26) \\ &= .9943 - .1038 \\ &= .8905 \end{aligned}$$

- b. According to the Rule of Thumb given in Section 5.4, n should be greater than 30 in order to apply the C.L.T., thus using the same procedure for $n = 15$ as was used for $n = 40$ would not be appropriate.

51.

$X \sim N(10, 4)$. For day 1, $n = 5$

$$P(\bar{X} \leq 11) = P\left(Z \leq \frac{11 - 10}{2/\sqrt{5}}\right) = P(Z \leq 1.12) = .8686$$

For day 2, $n = 6$

$$P(\bar{X} \leq 11) = P\left(Z \leq \frac{11 - 10}{2/\sqrt{6}}\right) = P(Z \leq 1.22) = .8888$$

For both days,

$$P(\bar{X} \leq 11) = (.8686)(.8888) = .7720$$

52.

$X \sim N(10)$, $n = 4$

$$\mu_{T_0} = n\mu = (4)(10) = 40 \text{ and } \sigma_{T_0} = \sigma\sqrt{n} = (2)(2) = 2,$$

We desire the 95th percentile: $40 + (1.645)(2) = 43.29$

Chapter 5: Joint Probability Distributions and Random Samples

53. $\mu = 50, \sigma = 1.2$

a. $n = 9$

$$P(\bar{X} \geq 51) = P\left(Z \geq \frac{51 - 50}{1.2/\sqrt{9}}\right) = P(Z \geq 2.5) = 1 - .9938 = .0062$$

b. $n = 40$

$$P(\bar{X} \geq 51) = P\left(Z \geq \frac{51 - 50}{1.2/\sqrt{40}}\right) = P(Z \geq 5.27) \approx 0$$

54.

a. $\mu_{\bar{X}} = \mu = 2.65, \sigma_{\bar{X}} = \frac{\sigma_x}{\sqrt{n}} = \frac{.85}{5} = .17$

$$P(\bar{X} \leq 3.00) = P\left(Z \leq \frac{3.00 - 2.65}{.17}\right) = P(Z \leq 2.06) = .9803$$

$$P(2.65 \leq \bar{X} \leq 3.00) = P(\bar{X} \leq 3.00) - P(\bar{X} \leq 2.65) = .4803$$

b. $P(\bar{X} \leq 3.00) = P\left(Z \leq \frac{3.00 - 2.65}{.85/\sqrt{n}}\right) = .99$ implies that $\frac{.35}{.85/\sqrt{n}} = 2.33$, from which $n = 32.02$. Thus $n = 33$ will suffice.

55. $\mu = np = 20 \quad \sigma = \sqrt{npq} = 3.464$

a. $P(25 \leq X) \approx P\left(\frac{24.5 - 20}{3.464} \leq Z\right) = P(1.30 \leq Z) = .0968$

b. $P(15 \leq X \leq 25) \approx P\left(\frac{14.5 - 20}{3.464} \leq Z \leq \frac{25.5 - 20}{3.464}\right)$
 $= P(-1.59 \leq Z \leq 1.59) = .8882$

56.

a. With $Y = \#$ of tickets, Y has approximately a normal distribution with $\mu = \lambda = 50$,

$$\sigma = \sqrt{\lambda} = 7.071, \text{ so } P(35 \leq Y \leq 70) \approx P\left(\frac{34.5 - 50}{7.071} \leq Z \leq \frac{70.5 - 50}{7.071}\right) = P(-2.19 \leq Z \leq 2.90) = .9838$$

b. Here $\mu = 250, \sigma^2 = 250, \sigma = 15.811$, so $P(225 \leq Y \leq 275) \approx$

$$P\left(\frac{224.5 - 250}{15.811} \leq Z \leq \frac{275.5 - 250}{15.811}\right) = P(-1.61 \leq Z \leq 1.61) = .8926$$

57. $E(X) = 100$, $\text{Var}(X) = 200$, $\sigma_x = 14.14$, so $P(X \leq 125) \approx P\left(Z \leq \frac{125 - 100}{14.14}\right)$
 $= P(Z \leq 1.77) = .9616$

Section 5.5

58.

- a. $E(27X_1 + 125X_2 + 512X_3) = 27E(X_1) + 125E(X_2) + 512E(X_3)$
 $= 27(200) + 125(250) + 512(100) = 87,850$
 $V(27X_1 + 125X_2 + 512X_3) = 27^2 V(X_1) + 125^2 V(X_2) + 512^2 V(X_3)$
 $= 27^2 (10)^2 + 125^2 (12)^2 + 512^2 (8)^2 = 19,100,116$
- b. The expected value is still correct, but the variance is not because the covariances now also contribute to the variance.

59.

- a. $E(X_1 + X_2 + X_3) = 180$, $V(X_1 + X_2 + X_3) = 45$, $\sigma_{X_1+X_2+X_3} = 6.708$
 $P(X_1 + X_2 + X_3 \leq 200) = P\left(Z \leq \frac{200 - 180}{6.708}\right) = P(Z \leq 2.98) = .9986$
 $P(150 \leq X_1 + X_2 + X_3 \leq 200) = P(-4.47 \leq Z \leq 2.98) \approx .9986$
- b. $\mu_{\bar{X}} = \mu = 60$, $\sigma_{\bar{X}} = \frac{\sigma_x}{\sqrt{n}} = \frac{\sqrt{15}}{\sqrt{3}} = 2.236$
 $P(\bar{X} \geq 55) = P\left(Z \geq \frac{55 - 60}{2.236}\right) = P(Z \geq -2.236) = .9875$
 $P(58 \leq \bar{X} \leq 62) = P(-.89 \leq Z \leq .89) = .6266$
- c. $E(X_1 - .5X_2 - .5X_3) = 0$;
 $V(X_1 - .5X_2 - .5X_3) = \sigma_1^2 + .25\sigma_2^2 + .25\sigma_3^2 = 22.5$, $\text{sd} = 4.7434$
 $P(-10 \leq X_1 - .5X_2 - .5X_3 \leq 5) = P\left(\frac{-10 - 0}{4.7434} \leq Z \leq \frac{5 - 0}{4.7434}\right)$
 $= P(-2.11 \leq Z \leq 1.05) = .8531 - .0174 = .8357$

Chapter 5: Joint Probability Distributions and Random Samples

d. $E(X_1 + X_2 + X_3) = 150$, $V(X_1 + X_2 + X_3) = 36$, $\sigma_{X_1+X_2+X_3} = 6$

$$P(X_1 + X_2 + X_3 \leq 200) = P\left(Z \leq \frac{160-150}{6}\right) = P(Z \leq 1.67) = .9525$$

We want $P(X_1 + X_2 \geq 2X_3)$, or written another way, $P(X_1 + X_2 - 2X_3 \geq 0)$

$$E(X_1 + X_2 - 2X_3) = 40 + 50 - 2(60) = -30,$$

$$V(X_1 + X_2 - 2X_3) = \sigma_1^2 + \sigma_2^2 + 4\sigma_3^2 = 78, 36, \text{sd} = 8.832, \text{so}$$

$$P(X_1 + X_2 - 2X_3 \geq 0) = P\left(Z \geq \frac{0 - (-30)}{8.832}\right) = P(Z \geq 3.40) = .0003$$

60. Y is normally distributed with $\mu_Y = \frac{1}{2}(\mu_1 + \mu_2) - \frac{1}{3}(\mu_3 + \mu_4 + \mu_5) = -1$, and

$$\sigma_Y^2 = \frac{1}{4}\sigma_1^2 + \frac{1}{4}\sigma_2^2 + \frac{1}{9}\sigma_3^2 + \frac{1}{9}\sigma_4^2 + \frac{1}{9}\sigma_5^2 = 3.167, \sigma_Y = 1.7795.$$

Thus, $P(0 \leq Y) = P\left(\frac{0 - (-1)}{1.7795} \leq Z\right) = P(.56 \leq Z) = .2877$ and

$$P(-1 \leq Y \leq 1) = P\left(0 \leq Z \leq \frac{2}{1.7795}\right) = P(0 \leq Z \leq 1.12) = .3686$$

61.

a. The marginal pmf's of X and Y are given in the solution to Exercise 7, from which $E(X) = 2.8$, $E(Y) = .7$, $V(X) = 1.66$, $V(Y) = .61$. Thus $E(X+Y) = E(X) + E(Y) = 3.5$, $V(X+Y) = V(X) + V(Y) = 2.27$, and the standard deviation of $X + Y$ is 1.51

b. $E(3X+10Y) = 3E(X) + 10E(Y) = 15.4$, $V(3X+10Y) = 9V(X) + 100V(Y) = 75.94$, and the standard deviation of revenue is 8.71

62. $E(X_1 + X_2 + X_3) = E(X_1) + E(X_2) + E(X_3) = 15 + 30 + 20 = 65$ min.,

$$V(X_1 + X_2 + X_3) = 1^2 + 2^2 + 1.5^2 = 7.25, \sigma_{X_1+X_2+X_3} = \sqrt{7.25} = 2.6926$$

Thus, $P(X_1 + X_2 + X_3 \leq 60) = P\left(Z \leq \frac{60-65}{2.6926}\right) = P(Z \leq -1.86) = .0314$

63.

a. $E(X_1) = 1.70$, $E(X_2) = 1.55$, $E(X_1X_2) = \sum_{x_1} \sum_{x_2} x_1x_2 p(x_1, x_2) = 3.33$, so $\text{Cov}(X_1, X_2) =$

$$E(X_1X_2) - E(X_1)E(X_2) = 3.33 - 2.635 = .695$$

b. $V(X_1 + X_2) = V(X_1) + V(X_2) + 2\text{Cov}(X_1, X_2)$
 $= 1.59 + 1.0875 + 2(.695) = 4.0675$

Chapter 5: Joint Probability Distributions and Random Samples

- 64.** Let $X_1, \dots, X_5$ denote morning times and $X_6, \dots, X_{10}$ denote evening times.
- a. $E(X_1 + \dots + X_{10}) = E(X_1) + \dots + E(X_{10}) = 5 E(X_1) + 5 E(X_6)$
 $= 5(4) + 5(5) = 45$
 - b. $\text{Var}(X_1 + \dots + X_{10}) = \text{Var}(X_1) + \dots + \text{Var}(X_{10}) = 5 \text{Var}(X_1) + 5 \text{Var}(X_6)$
 $= 5 \left[\frac{64}{12} + \frac{100}{12} \right] = \frac{820}{12} = 68.33$
 - c. $E(X_1 - X_6) = E(X_1) - E(X_6) = 4 - 5 = -1$
 $\text{Var}(X_1 - X_6) = \text{Var}(X_1) + \text{Var}(X_6) = \frac{64}{12} + \frac{100}{12} = \frac{164}{12} = 13.67$
 - d. $E[(X_1 + \dots + X_5) - (X_6 + \dots + X_{10})] = 5(4) - 5(5) = -5$
 $\text{Var}[(X_1 + \dots + X_5) - (X_6 + \dots + X_{10})]$
 $= \text{Var}(X_1 + \dots + X_5) + \text{Var}(X_6 + \dots + X_{10}) = 68.33$
- 65.** $\mu = 5.00, \sigma = .2$
- a. $E(\bar{X} - \bar{Y}) = 0; \quad V(\bar{X} - \bar{Y}) = \frac{\sigma^2}{25} + \frac{\sigma^2}{25} = .0032, \sigma_{\bar{X} - \bar{Y}} = .0566$
 $\Rightarrow P(-.1 \leq \bar{X} - \bar{Y} \leq .1) \approx P(-1.77 \leq Z \leq 1.77) = .9232$ (by the CLT)
 - b. $V(\bar{X} - \bar{Y}) = \frac{\sigma^2}{36} + \frac{\sigma^2}{36} = .0022222, \sigma_{\bar{X} - \bar{Y}} = .0471$
 $\Rightarrow P(-.1 \leq \bar{X} - \bar{Y} \leq .1) \approx P(-2.12 \leq Z \leq 2.12) = .9660$
- 66.**
- a. With $M = 5X_1 + 10X_2$, $E(M) = 5(2) + 10(4) = 50$,
 $\text{Var}(M) = 5^2 (.5)^2 + 10^2 (1)^2 = 106.25, \sigma_M = 10.308$.
 - b. $P(75 < M) = P\left(\frac{75 - 50}{10.308} < Z\right) = P(2.43 < Z) = .0075$
 - c. $M = A_1X_1 + A_2X_2$ with the A_i 's and X_i 's all independent, so
 $E(M) = E(A_1X_1) + E(A_2X_2) = E(A_1)E(X_1) + E(A_2)E(X_2) = 50$
 - d. $\text{Var}(M) = E(M^2) - [E(M)]^2$. Recall that for any r.v. Y ,
 $E(Y^2) = \text{Var}(Y) + [E(Y)]^2$. Thus, $E(M^2) = E(A_1^2 X_1^2 + 2A_1 X_1 A_2 X_2 + A_2^2 X_2^2)$
 $= E(A_1^2)E(X_1^2) + 2E(A_1)E(X_1)E(A_2)E(X_2) + E(A_2^2)E(X_2^2)$
 (by independence)
 $= (.25 + 25)(.25 + 4) + 2(5)(2)(10)(4) + (.25 + 100)(1 + 16) = 2611.5625$, so $\text{Var}(M) = 2611.5625 - (50)^2 = 111.5625$

Chapter 5: Joint Probability Distributions and Random Samples

- e. $E(M) = 50$ still, but now

$$\begin{aligned} \text{Var}(M) &= a_1^2 \text{Var}(X_1) + 2a_1 a_2 \text{Cov}(X_1, X_2) + a_2^2 \text{Var}(X_2) \\ &= 6.25 + 2(5)(10)(-.25) + 100 = 81.25 \end{aligned}$$

67. Letting X_1 , X_2 , and X_3 denote the lengths of the three pieces, the total length is $X_1 + X_2 - X_3$. This has a normal distribution with mean value $20 + 15 - 1 = 34$, variance $.25 + .16 + .01 = .42$, and standard deviation $.6481$. Standardizing gives
 $P(34.5 \leq X_1 + X_2 - X_3 \leq 35) = P(.77 \leq Z \leq 1.54) = .1588$

68. Let X_1 and X_2 denote the (constant) speeds of the two planes.

- a. After two hours, the planes have traveled $2X_1$ km. and $2X_2$ km., respectively, so the second will not have caught the first if $2X_1 + 10 > 2X_2$, i.e. if $X_2 - X_1 < 5$. $X_2 - X_1$ has a mean $500 - 520 = -20$, variance $100 + 100 = 200$, and standard deviation 14.14 . Thus,

$$P(X_2 - X_1 < 5) = P\left(Z < \frac{5 - (-20)}{14.14}\right) = P(Z < 1.77) = .9616.$$

- b. After two hours, #1 will be $10 + 2X_1$ km from where #2 started, whereas #2 will be $2X_2$ from where it started. Thus the separation distance will be at most 10 if $|2X_2 - 10 - 2X_1| \leq 10$, i.e. $-10 \leq 2X_2 - 10 - 2X_1 \leq 10$, i.e. $0 \leq X_2 - X_1 \leq 10$. The corresponding probability is
 $P(0 \leq X_2 - X_1 \leq 10) = P(1.41 \leq Z \leq 2.12) = .9830 - .9207 = .0623$.

- 69.

- a. $E(X_1 + X_2 + X_3) = 800 + 1000 + 600 = 2400$.
- b. Assuming independence of X_1 , X_2 , X_3 , $\text{Var}(X_1 + X_2 + X_3) = (16)^2 + (25)^2 + (18)^2 = 12.05$
- c. $E(X_1 + X_2 + X_3) = 2400$ as before, but now $\text{Var}(X_1 + X_2 + X_3) = \text{Var}(X_1) + \text{Var}(X_2) + \text{Var}(X_3) + 2\text{Cov}(X_1, X_2) + 2\text{Cov}(X_1, X_3) + 2\text{Cov}(X_2, X_3) = 1745$, with $\text{sd} = 41.77$

- 70.

- a. $E(Y_i) = .5$, so $E(W) = \sum_{i=1}^n i \cdot E(Y_i) = .5 \sum_{i=1}^n i = \frac{n(n+1)}{4}$
- b. $\text{Var}(Y_i) = .25$, so $\text{Var}(W) = \sum_{i=1}^n i^2 \cdot \text{Var}(Y_i) = .25 \sum_{i=1}^n i^2 = \frac{n(n+1)(2n+1)}{24}$

71.

- a. $M = a_1 X_1 + a_2 X_2 + W \int_0^{12} x dx = a_1 X_1 + a_2 X_2 + 72W$, so
 $E(M) = (5)(2) + (10)(4) + (72)(1.5) = 158m$
 $\sigma_M^2 = (5)^2(.5)^2 + (10)^2(1)^2 + (72)^2(.25)^2 = 430.25$, $\sigma_M = 20.74$
- b. $P(M \leq 200) = P\left(Z \leq \frac{200-158}{20.74}\right) = P(Z \leq 2.03) = .9788$

72.

The total elapsed time between leaving and returning is $T_o = X_1 + X_2 + X_3 + X_4$, with
 $E(T_o) = 40$, $\sigma_{T_o}^2 = 40$, $\sigma_{T_o} = 5.477$. T_o is normally distributed, and the desired value t
 is the 99th percentile of the lapsed time distribution added to 10 A.M.: 10:00 +
 $[40+(5.477)(2.33)] = 10:52.76$

73.

- a. Both approximately normal by the C.L.T.
- b. The difference of two r.v.'s is just a special linear combination, and a linear combination of normal r.v.'s has a normal distribution, so $\bar{X} - \bar{Y}$ has approximately a normal distribution with $\mu_{\bar{X}-\bar{Y}} = 5$ and $\sigma_{\bar{X}-\bar{Y}}^2 = \frac{8^2}{40} + \frac{6^2}{35} = 2.629$, $\sigma_{\bar{X}-\bar{Y}} = 1.621$
- c. $P(-1 \leq \bar{X} - \bar{Y} \leq 1) \approx P\left(\frac{-1-5}{1.6213} \leq Z \leq \frac{1-5}{1.6213}\right)$
 $= P(-3.70 \leq Z \leq -2.47) \approx .0068$
- d. $P(\bar{X} - \bar{Y} \geq 10) \approx P\left(Z \geq \frac{10-5}{1.6213}\right) = P(Z \geq 3.08) = .0010$. This probability is quite small, so such an occurrence is unlikely if $\mu_1 - \mu_2 = 5$, and we would thus doubt this claim.

74.

X is approximately normal with $\mu_1 = (50)(.7) = 35$ and $\sigma_1^2 = (50)(.7)(.3) = 10.5$, as is Y with $\mu_2 = 30$ and $\sigma_2^2 = 12$. Thus $\mu_{X-Y} = 5$ and $\sigma_{X-Y}^2 = 22.5$, so
 $p(-5 \leq X - Y \leq 5) \approx P\left(\frac{-10}{4.74} \leq Z \leq \frac{0}{4.74}\right) = P(-2.11 \leq Z \leq 0) = .4826$

Supplementary Exercises

75.

- a. $p_X(x)$ is obtained by adding joint probabilities across the row labeled x , resulting in $p_X(x) = .2, .5, .3$ for $x = 12, 15, 20$ respectively. Similarly, from column sums $p_Y(y) = .1, .35, .55$ for $y = 12, 15, 20$ respectively.
- b. $P(X \leq 15 \text{ and } Y \leq 15) = p(12,12) + p(12,15) + p(15,12) + p(15,15) = .25$
- c. $p_X(12) \cdot p_Y(12) = (.2)(.1) \neq .05 = p(12,12)$, so X and Y are not independent. (Almost any other (x,y) pair yields the same conclusion).
- d. $E(X + Y) = \sum \sum (x + y) p(x, y) = 33.35$ (or $= E(X) + E(Y) = 33.35$)
- e. $E(|X - Y|) = \sum \sum |x - y| p(x, y) = 3.85$

76. The roll-up procedure is not valid for the 75th percentile unless $\sigma_1 = 0$ or $\sigma_2 = 0$ or both σ_1 and $\sigma_2 = 0$, as described below.

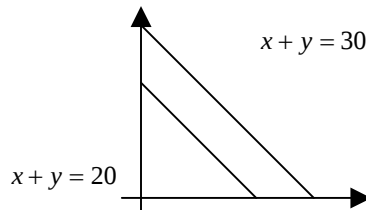
$$\text{Sum of percentiles: } \mu_1 + (Z)\sigma_1 + \mu_2 + (Z)\sigma_2 = \mu_1 + \mu_2 + (Z)(\sigma_1 + \sigma_2)$$

$$\text{Percentile of sums: } \mu_1 + \mu_2 + (Z)\sqrt{\sigma_1^2 + \sigma_2^2}$$

These are equal when $Z = 0$ (i.e. for the median) or in the unusual case when

$$\sigma_1 + \sigma_2 = \sqrt{\sigma_1^2 + \sigma_2^2}, \text{ which happens when } \sigma_1 = 0 \text{ or } \sigma_2 = 0 \text{ or both } \sigma_1 \text{ and } \sigma_2 = 0.$$

77.



$$\begin{aligned} \text{a. } 1 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = \int_0^{20} \int_{20-x}^{30-x} kxy dy dx + \int_{20}^{30} \int_0^{30-x} kxy dy dx \\ &= \frac{81,250}{3} \cdot k \Rightarrow k = \frac{3}{81,250} \end{aligned}$$

$$\text{b. } f_X(x) = \begin{cases} \int_{20-x}^{30-x} kxy dy = k(250x - 10x^2) & 0 \leq x \leq 20 \\ \int_0^{30-x} kxy dy = k(450x - 30x^2 + \frac{1}{2}x^3) & 20 \leq x \leq 30 \end{cases}$$

and by symmetry $f_Y(y)$ is obtained by substituting y for x in $f_X(x)$. Since $f_X(25) > 0$, and $f_Y(25) > 0$, but $f(25, 25) = 0$, $f_X(x) \cdot f_Y(y) \neq f(x, y)$ for all x, y so X and Y are not independent.

$$\begin{aligned} \text{c. } P(X + Y \leq 25) &= \int_0^{20} \int_{20-x}^{25-x} kxy dy dx + \int_{20}^{25} \int_0^{25-x} kxy dy dx \\ &= \frac{3}{81,250} \cdot \frac{230,625}{24} = .355 \end{aligned}$$

$$\begin{aligned} \text{d. } E(X + Y) &= E(X) + E(Y) = 2 \left\{ \int_0^{20} x \cdot k(250x - 10x^2) dx \right. \\ &\quad \left. + \int_{20}^{30} x \cdot k(450x - 30x^2 + \frac{1}{2}x^3) dx \right\} = 2k(351,666.67) = 25.969 \end{aligned}$$

$$\begin{aligned} \text{e. } E(XY) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy \cdot f(x, y) dx dy = \int_0^{20} \int_{20-x}^{30-x} kx^2 y^2 dy dx \\ &\quad + \int_{20}^{30} \int_0^{30-x} kx^2 y^2 dy dx = \frac{k}{3} \cdot \frac{33,250,000}{3} = 136.4103, \text{ so} \\ \text{Cov}(X, Y) &= 136.4103 - (12.9845)^2 = -32.19, \text{ and } E(X^2) = E(Y^2) = 204.6154, \text{ so} \\ \sigma_x^2 &= \sigma_y^2 = 204.6154 - (12.9845)^2 = 36.0182 \text{ and } \rho = \frac{-32.19}{36.0182} = -.894 \end{aligned}$$

$$\text{f. } \text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y) = 7.66$$

$$\begin{aligned} 78. \quad F_Y(y) &= P(\max(X_1, \dots, X_n) \leq y) = P(X_1 \leq y, \dots, X_n \leq y) = [P(X_1 \leq y)]^n = \left(\frac{y-100}{100} \right)^n \text{ for} \\ &100 \leq y \leq 200. \end{aligned}$$

$$\text{Thus } f_Y(y) = \frac{n}{100^n} (y-100)^{n-1} \text{ for } 100 \leq y \leq 200.$$

$$\begin{aligned} E(Y) &= \int_{100}^{200} y \cdot \frac{n}{100^n} (y-100)^{n-1} dy = \frac{n}{100^n} \int_0^{100} (u+100) u^{n-1} du \\ &= 100 + \frac{n}{100^n} \int_0^{100} u^n du = 100 + 100 \frac{n}{n+1} = \frac{2n+1}{n+1} \cdot 100 \end{aligned}$$

$$79. \quad E(\bar{X} + \bar{Y} + \bar{Z}) = 500 + 900 + 2000 = 3400$$

$$\text{Var}(\bar{X} + \bar{Y} + \bar{Z}) = \frac{50^2}{365} + \frac{100^2}{365} + \frac{180^2}{365} = 123.014, \text{ and the std dev} = 11.09.$$

$$P(\bar{X} + \bar{Y} + \bar{Z} \leq 3500) = P(Z \leq 9.0) \approx 1$$

Chapter 5: Joint Probability Distributions and Random Samples

80.

- a. Let $X_1, \dots, X_{12}$ denote the weights for the business-class passengers and $Y_1, \dots, Y_{50}$ denote the tourist-class weights. Then T = total weight
 $= X_1 + \dots + X_{12} + Y_1 + \dots + Y_{50} = X + Y$
 $E(X) = 12E(X_1) = 12(30) = 360$; $V(X) = 12V(X_1) = 12(36) = 432$.
 $E(Y) = 50E(Y_1) = 50(40) = 2000$; $V(Y) = 50V(Y_1) = 50(100) = 5000$.
 Thus $E(T) = E(X) + E(Y) = 360 + 2000 = 2360$
 And $V(T) = V(X) + V(Y) = 432 + 5000 = 5432$, std dev = 73.7021

b. $P(T \leq 2500) = P\left(Z \leq \frac{2500 - 2360}{73.7021}\right) = P(Z \leq 1.90) = .9713$

81.

- a. $E(N) \cdot \mu = (10)(40) = 400$ minutes
- b. We expect 20 components to come in for repair during a 4 hour period,
 so $E(N) \cdot \mu = (20)(3.5) = 70$

82. $X \sim \text{Bin}(200, .45)$ and $Y \sim \text{Bin}(300, .6)$. Because both n 's are large, both X and Y are approximately normal, so $X + Y$ is approximately normal with mean $(200)(.45) + (300)(.6) = 270$, variance $200(.45)(.55) + 300(.6)(.4) = 121.40$, and standard deviation 11.02. Thus, $P(X$

$+ Y \geq 250) = P\left(Z \geq \frac{249.5 - 270}{11.02}\right) = P(Z \geq -1.86) = .9686$

83. $0.95 = P(\mu - .02 \leq \bar{X} \leq \mu + .02) = P\left(\frac{-.02}{.01/\sqrt{n}} \leq Z \leq \frac{.02}{.01/\sqrt{n}}\right)$
 $= P(-.2\sqrt{n} \leq Z \leq .2\sqrt{n})$, but $P(-1.96 \leq Z \leq 1.96) = .95$ so
 $.2\sqrt{n} = 1.96 \Rightarrow n = 97$. The C.L.T.

84. I have 192 oz. The amount which I would consume if there were no limit is $T_0 = X_1 + \dots + X_{14}$ where each X_i is normally distributed with $\mu = 13$ and $\sigma = 2$. Thus T_0 is normal with $\mu_{T_0} = 182$ and $\sigma_{T_0} = 7.483$, so $P(T_0 < 192) = P(Z < 1.34) = .9099$.

85. The expected value and standard deviation of volume are 87,850 and 4370.37, respectively, so

$P(\text{volume} \leq 100,000) = P\left(Z \leq \frac{100,000 - 87,850}{4370.37}\right) = P(Z \leq 2.78) = .9973$

86. The student will not be late if $X_1 + X_3 \leq X_2$, i.e. if $X_1 - X_2 + X_3 \leq 0$. This linear combination has mean -2 , variance 4.25, and standard deviation 2.06, so

$P(X_1 - X_2 + X_3 \leq 0) = P\left(Z \leq \frac{0 - (-2)}{2.06}\right) = P(Z \leq .97) = .8340$

87.

a. $Var(aX + Y) = a^2 \sigma_x^2 + 2aCov(X, Y) + \sigma_y^2 = a^2 \sigma_x^2 + 2a\sigma_x \sigma_y \rho + \sigma_y^2.$

Substituting $a = \frac{\sigma_y}{\sigma_x}$ yields $\sigma_y^2 + 2\sigma_y^2 \rho + \sigma_y^2 = 2\sigma_y^2(1 - \rho) \geq 0$, so $\rho \geq -1$

b. Same argument as in a

c. Suppose $\rho = 1$. Then $Var(aX - Y) = 2\sigma_y^2(1 - \rho) = 0$, which implies that $aX - Y = k$ (a constant), so $aX - Y = aX - k$, which is of the form $aX + b$.

88. $E(X + Y - t)^2 = \int_0^1 \int_0^1 (x + y - t)^2 \cdot f(x, y) dx dy$. To find the minimizing value of t , take the derivative with respect to t and equate it to 0:

$$0 = \int_0^1 \int_0^1 2(x + y - t)(-1) f(x, y) dx dy = 0 \Rightarrow \int_0^1 \int_0^1 t f(x, y) dx dy = t$$

$$= \int_0^1 \int_0^1 (x + y) \cdot f(x, y) dx dy = E(X + Y), \text{ so the best prediction is the individual's expected score } (= 1.167).$$

89.

a. With $Y = X_1 + X_2$,

$$F_Y(y) = \int_0^y \left\{ \int_0^{y-x_1} \frac{1}{2^{\nu_1/2} \Gamma(\nu_1/2)} \cdot \frac{1}{2^{\nu_2/2} \Gamma(\nu_2/2)} \cdot x_1^{\frac{\nu_1}{2}-1} x_2^{\frac{\nu_2}{2}-1} e^{-\frac{x_1+x_2}{2}} dx_2 \right\} dx_1.$$

But the inner integral can be shown to be equal to

$$\frac{1}{2^{(\nu_1+\nu_2)/2} \Gamma((\nu_1 + \nu_2)/2)} y^{[(\nu_1+\nu_2)/2]-1} e^{-y/2}, \text{ from which the result follows.}$$

b. By a, $Z_1^2 + Z_2^2$ is chi-squared with $\nu = 2$, so $(Z_1^2 + Z_2^2) + Z_3^2$ is chi-squared with $\nu = 3$, etc, until $Z_1^2 + \dots + Z_n^2$ is chi-squared with $\nu = n$

c. $\frac{X_i - \mu}{\sigma}$ is standard normal, so $\left[\frac{X_i - \mu}{\sigma} \right]^2$ is chi-squared with $\nu = 1$, so the sum is chi-squared with $\nu = n$.

90.

- a. $\begin{aligned}\text{Cov}(X, Y + Z) &= E[X(Y + Z)] - E(X) \cdot E(Y + Z) \\ &= E(XY) + E(XZ) - E(X) \cdot E(Y) - E(X) \cdot E(Z) \\ &= E(XY) - E(X) \cdot E(Y) + E(XZ) - E(X) \cdot E(Z) \\ &= \text{Cov}(X, Y) + \text{Cov}(X, Z).\end{aligned}$
- b. $\text{Cov}(X_1 + X_2, Y_1 + Y_2) = \text{Cov}(X_1, Y_1) + \text{Cov}(X_1, Y_2) + \text{Cov}(X_2, Y_1) + \text{Cov}(X_2, Y_2)$
(apply **a** twice) = 16.

91.

- a. $V(X_1) = V(W + E_1) = \sigma_W^2 + \sigma_E^2 = V(W + E_2) = V(X_2)$ and
 $\text{Cov}(X_1, X_2) = \text{Cov}(W + E_1, W + E_2) = \text{Cov}(W, W) + \text{Cov}(W, E_2) + \text{Cov}(E_1, W) + \text{Cov}(E_1, E_2) = \text{Cov}(W, W) = V(W) = \sigma_W^2$.

$$\text{Thus, } \rho = \frac{\sigma_W^2}{\sqrt{\sigma_W^2 + \sigma_E^2} \cdot \sqrt{\sigma_W^2 + \sigma_E^2}} = \frac{\sigma_W^2}{\sigma_W^2 + \sigma_E^2}$$

- b. $\rho = \frac{1}{1 + .0001} = .9999$

92.

- a. $\begin{aligned}\text{Cov}(X, Y) &= \text{Cov}(A + D, B + E) \\ &= \text{Cov}(A, B) + \text{Cov}(D, B) + \text{Cov}(A, E) + \text{Cov}(D, E) = \text{Cov}(A, B).\end{aligned}$ Thus

$$\begin{aligned}\text{Corr}(X, Y) &= \frac{\text{Cov}(A, B)}{\sqrt{\sigma_A^2 + \sigma_D^2} \cdot \sqrt{\sigma_B^2 + \sigma_E^2}} \\ &= \frac{\text{Cov}(A, B)}{\sigma_A \sigma_B} \cdot \frac{\sigma_A}{\sqrt{\sigma_A^2 + \sigma_D^2}} \cdot \frac{\sigma_B}{\sqrt{\sigma_B^2 + \sigma_E^2}}\end{aligned}$$

The first factor in this expression is $\text{Corr}(A, B)$, and (by the result of exercise 70a) the second and third factors are the square roots of $\text{Corr}(X_1, X_2)$ and $\text{Corr}(Y_1, Y_2)$, respectively. Clearly, measurement error reduces the correlation, since both square-root factors are between 0 and 1.

- b. $\sqrt{.8100} \cdot \sqrt{.9025} = .855$. This is disturbing, because measurement error substantially reduces the correlation.

Chapter 5: Joint Probability Distributions and Random Samples

93. $E(Y) \doteq h(\mu_1, \mu_2, \mu_3, \mu_4) = 120\left[\frac{1}{10} + \frac{1}{15} + \frac{1}{20}\right] = 26$

The partial derivatives of $h(\mu_1, \mu_2, \mu_3, \mu_4)$ with respect to x_1, x_2, x_3 , and x_4 are $-\frac{x_4}{x_1^2}$,

$-\frac{x_4}{x_2^2}$, $-\frac{x_4}{x_3^2}$, and $\frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3}$, respectively. Substituting $x_1 = 10, x_2 = 15, x_3 = 20$, and

$x_4 = 120$ gives $-1.2, -.5333, -.3000$, and $.2167$, respectively, so $V(Y) = (1)(-1.2)^2 + (1)(-.5333)^2 + (1.5)(-.3000)^2 + (4.0)(.2167)^2 = 2.6783$, and the approximate sd of y is 1.64 .

94. The four second order partials are $\frac{2x_4}{x_1^3}, \frac{2x_4}{x_2^3}, \frac{2x_4}{x_3^3}$, and 0 respectively. Substitution gives

$$E(Y) = 26 + .1200 + .0356 + .0338 = 26.1894.$$

CHAPTER 7

Section 7.1

1.

- a. $z_{\alpha/2} = 2.81$ implies that $\alpha/2 = 1 - \Phi(2.81) = .0025$, so $\alpha = .005$ and the confidence level is $100(1 - \alpha)\% = 99.5\%$.
- b. $z_{\alpha/2} = 1.44$ for $\alpha = 2[1 - \Phi(1.44)] = .15$, and $100(1 - \alpha)\% = 85\%$.
- c. 99.7% implies that $\alpha = .003$, $\alpha/2 = .0015$, and $z_{.0015} = 2.96$. (Look for cumulative area .9985 in the main body of table A.3, the Z table.)
- d. 75% implies $\alpha = .25$, $\alpha/2 = .125$, and $z_{.125} = 1.15$.

2.

- a. The sample mean is the center of the interval, so $\bar{x} = \frac{114.4 + 115.6}{2} = 115$.
- b. The interval (114.4, 115.6) has the 90% confidence level. The higher confidence level will produce a wider interval.

3.

- a. A 90% confidence interval will be narrower (See 2b, above) Also, the z critical value for a 90% confidence level is 1.645, smaller than the z of 1.96 for the 95% confidence level, thus producing a narrower interval.
- b. Not a correct statement. Once an interval has been created from a sample, the mean μ is either enclosed by it, or not. The 95% confidence is in the general procedure, for repeated sampling.
- c. Not a correct statement. The interval is an estimate for the population mean, not a boundary for population values.
- d. Not a correct statement. In theory, if the process were repeated an infinite number of times, 95% of the intervals would contain the population mean μ .

Chapter 7: Statistical Intervals Based on a Single Sample

4.

a. $58.3 \pm \frac{1.96(3)}{\sqrt{25}} = 58.3 \pm 1.18 = (57.1, 59.5)$

b. $58.3 \pm \frac{1.96(3)}{\sqrt{100}} = 58.3 \pm .59 = (57.7, 58.9)$

c. $58.3 \pm \frac{2.58(3)}{\sqrt{100}} = 58.3 \pm .77 = (57.5, 59.1)$

d. 82% confidence $\Rightarrow 1 - \alpha = .82 \Rightarrow \alpha = .18 \Rightarrow \alpha/2 = .09$, so $z_{\alpha/2} = z_{.09} = 1.34$ and the interval is $58.3 \pm \frac{1.34(3)}{\sqrt{100}} = (57.9, 58.7)$.

e. $n = \left[\frac{2(2.58)3}{1} \right]^2 = 239.62$ so $n = 240$.

5.

a. $4.85 \pm \frac{(1.96)(.75)}{\sqrt{20}} = 4.85 \pm .33 = (4.52, 5.18)$.

b. $z_{\alpha/2} = z_{.02/2} = z_{.01} = 2.33$, so the interval is $4.56 \pm \frac{(2.33)(.75)}{\sqrt{16}} = (4.12, 5.00)$.

c. $n = \left[\frac{2(1.96)(.75)}{.40} \right]^2 = 54.02$, so $n = 55$.

d. $n = \left[\frac{2(2.58)(.75)}{.2} \right]^2 = 93.61$, so $n = 94$.

6.

a. $8439 \pm \frac{(1.645)(100)}{\sqrt{25}} = 8439 \pm 32.9 = (8406.1, 8471.9)$.

b. $1 - \alpha = .92 \Rightarrow \alpha = .08 \Rightarrow \alpha/2 = .04$ so $z_{\alpha/2} = z_{.04} = 1.75$

Chapter 7: Statistical Intervals Based on a Single Sample

7. If $L = 2z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$ and we increase the sample size by a factor of 4, the new length is
- $$L' = 2z_{\alpha/2} \frac{\sigma}{\sqrt{4n}} = \left[2z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right] \left(\frac{1}{2} \right) = \frac{L}{2}.$$
- Thus halving the length requires n to be increased fourfold. If $n' = 25n$, then $L' = \frac{L}{5}$, so the length is decreased by a factor of 5.

8.

- a. With probability $1 - \alpha$, $z_{\alpha_1} \leq (\bar{X} - \mu) \left(\frac{\sigma}{\sqrt{n}} \right) \leq z_{\alpha_2}$. These inequalities can be manipulated exactly as was done in the text to isolate μ ; the result is

$$\bar{X} - z_{\alpha_2} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + z_{\alpha_1} \frac{\sigma}{\sqrt{n}}, \text{ so a } 100(1 - \alpha)\% \text{ interval is}$$

$$\left(\bar{X} - z_{\alpha_2} \frac{\sigma}{\sqrt{n}}, \bar{X} + z_{\alpha_1} \frac{\sigma}{\sqrt{n}} \right)$$

- b. The usual 95% interval has length $3.92 \frac{\sigma}{\sqrt{n}}$, while this interval will have length

$$(z_{\alpha_1} + z_{\alpha_2}) \frac{\sigma}{\sqrt{n}}. \text{ With } z_{\alpha_1} = z_{.0125} = 2.24 \text{ and } z_{\alpha_2} = z_{.0375} = 1.78, \text{ the length is}$$

$$(2.24 + 1.78) \frac{\sigma}{\sqrt{n}} = 4.02 \frac{\sigma}{\sqrt{n}}, \text{ which is longer.}$$

9.

- a. $\left(\bar{x} - 1.645 \frac{\sigma}{\sqrt{n}}, \infty \right)$. From 5a, $\bar{x} = 4.85$, $\sigma = .75$, $n = 20$;
- $$4.85 - 1.645 \frac{.75}{\sqrt{20}} = 4.5741, \text{ so the interval is } (4.5741, \infty).$$

- b. $\left(\bar{x} - z_{\alpha} \frac{\sigma}{\sqrt{n}}, \infty \right)$

- c. $\left(-\infty, \bar{x} + z_{\alpha} \frac{\sigma}{\sqrt{n}} \right)$; From 4a, $\bar{x} = 58.3$, $\sigma = 3.0$, $n = 25$;
- $$58.3 + 2.33 \frac{3}{\sqrt{25}} = (-\infty, 59.70)$$

Chapter 7: Statistical Intervals Based on a Single Sample

10.

- a. When $n = 15$, $2\lambda \sum X_i$ has a chi-squared distribution with 30 d.f. From the 30 d.f. row of Table A.6, the critical values that capture lower and upper tail areas of .025 (and thus a central area of .95) are 16.791 and 46.979. An argument parallel to that given in

Example 7.5 gives $\left(\frac{2\sum x_i}{46.979}, \frac{2\sum x_i}{16.791} \right)$ as a 95% C. I. for $\mu = \frac{1}{\lambda}$. Since

$\sum x_i = 63.2$ the interval is (2.69, 7.53).

- b. A 99% confidence level requires using critical values that capture area .005 in each tail of the chi-squared curve with 30 d.f.; these are 13.787 and 53.672, which replace 16.791 and 46.979 in a.

- c. $V(X) = \frac{1}{\lambda^2}$ when X has an exponential distribution, so the standard deviation is $\frac{1}{\lambda}$, the same as the mean. Thus the interval of a is also a 95% C.I. for the standard deviation of the lifetime distribution.

11. Y is a binomial r.v. with $n = 1000$ and $p = .95$, so $E(Y) = np = 950$, the expected number of intervals that capture μ , and $\sigma_Y = \sqrt{npq} = 6.892$. Using the normal approximation to the binomial distribution, $P(940 \leq Y \leq 960) = P(939.5 \leq Y_{\text{normal}} \leq 960.5) = P(-1.52 \leq Z \leq 1.52) = .9357 - .0643 = .8714$.

Section 7.2

12. $\bar{x} \pm 2.58 \frac{s}{\sqrt{n}} = .81 \pm 2.58 \frac{.34}{\sqrt{110}} = .81 \pm .08 = (.73, .89)$

13.

a. $\bar{x} \pm z_{.025} \frac{s}{\sqrt{n}} = 1.028 \pm 1.96 \frac{.163}{\sqrt{69}} = 1.028 \pm .038 = (.990, 1.066)$

b. $w = .05 = \frac{2(1.96)(.16)}{\sqrt{n}} \Rightarrow \sqrt{n} = \frac{2(1.96)(.16)}{.05} = 12.544 \Rightarrow n = (12.544)^2 \approx 158$

Chapter 7: Statistical Intervals Based on a Single Sample

14.

- a. $89.10 \pm 1.96 \frac{3.73}{\sqrt{169}} = 89.10 \pm .56 = (88.54, 89.66)$. Yes, this is a very narrow interval. It appears quite precise.

b. $n = \left[\frac{(1.96)(.16)}{.5} \right]^2 = 245.86 \Rightarrow n = 246$.

15.

- a. $z_\alpha = .84$, and $\Phi(.84) = .7995 \approx .80$, so the confidence level is 80%.

- b. $z_\alpha = 2.05$, and $\Phi(2.05) = .9798 \approx .98$, so the confidence level is 98%.

- c. $z_\alpha = .67$, and $\Phi(.67) = .7486 \approx .75$, so the confidence level is 75%.

16. $n = 46$, $\bar{x} = 382.1$, $s = 31.5$; The 95% upper confidence bound =

$$\bar{x} + z_\alpha \frac{s}{\sqrt{n}} = 382.1 + 1.645 \frac{31.5}{\sqrt{46}} = 382.1 + 7.64 = 389.74$$

17. $\bar{x} - z_{.01} \frac{s}{\sqrt{n}} = 135.39 - 2.33 \frac{4.59}{\sqrt{153}} = 135.39 - .865 = 134.53$ With a confidence level of 99%, the true average ultimate tensile strength is between $(134.53, \infty)$.

18. 90% lower bound: $\bar{x} - z_{.10} \frac{s}{\sqrt{n}} = 4.25 - 1.28 \frac{1.30}{\sqrt{75}} = 4.06$

19. $\hat{p} = \frac{201}{356} = .5646$; We calculate a 95% confidence interval for the proportion of all dies

that pass the probe:

$$\frac{.5646 + \frac{(1.96)^2}{2(356)} \pm 1.96 \sqrt{\frac{(.5646)(.4354)}{356} + \frac{(1.96)^2}{4(356)^2}}}{1 + \frac{(1.96)^2}{356}} = \frac{.5700 \pm .0518}{1.01079} = (.513, .615)$$

Chapter 7: Statistical Intervals Based on a Single Sample

20. Because the sample size is so large, the simpler formula (7.11) for the confidence interval for p is sufficient.

$$.15 \pm 2.58 \sqrt{\frac{(.15)(.85)}{4722}} = .15 \pm .013 = (.137, .163)$$

21. $\hat{p} = \frac{133}{539} = .2468$; the 95% lower confidence bound is:

$$\frac{.2468 + \frac{(1.645)^2}{2(539)} - 1.645 \sqrt{\frac{(.2468)(.7532)}{539} + \frac{(1.645)^2}{4(539)^2}}}{1 + \frac{(1.645)^2}{539}} = \frac{.2493 - .0307}{1.005} = .218$$

22. $\hat{p} = .072$; the 99% upper confidence bound is:

$$\frac{.072 + \frac{(2.33)^2}{2(487)} + 2.33 \sqrt{\frac{(.072)(.928)}{487} + \frac{(2.33)^2}{4(487)^2}}}{1 + \frac{(2.33)^2}{487}} = \frac{.0776 + .0279}{1.0111} = .1043$$

- 23.

- a. $\hat{p} = \frac{24}{37} = .6486$; The 99% confidence interval for p is

$$\frac{.6486 + \frac{(2.58)^2}{2(37)} \pm 2.58 \sqrt{\frac{(.6486)(.3514)}{37} + \frac{(2.58)^2}{4(37)^2}}}{1 + \frac{(2.58)^2}{37}} = \frac{.7386 \pm .2216}{1.1799} = (.438, .814)$$

- b.
$$n = \frac{2(2.58)^2(.25) - (2.58)^2(.01) \pm \sqrt{4(2.58)^4(.25)(.25 - .01) + .01(2.58)^4}}{.01}$$

$$= \frac{3.261636 \pm 3.3282}{.01} \approx 659$$

24. $n = 56$, $\bar{x} = 8.17$, $s = 1.42$; For a 95% C.I., $z_{\alpha/2} = 1.96$. The interval is

$$8.17 \pm 1.96 \left(\frac{1.42}{\sqrt{56}} \right) = (7.798, 8.542). \text{ We make no assumptions about the distribution if percentage elongation.}$$

Chapter 7: Statistical Intervals Based on a Single Sample

25.

$$\text{a. } n = \frac{2(1.96)^2(.25) - (1.96)^2(.01) \pm \sqrt{4(1.96)^4(.25)(.25 - .01) + .01(1.96)^4}}{.01} \approx 381$$

$$\text{b. } n = \frac{2(1.96)^2\left(\frac{1}{3} \cdot \frac{2}{3}\right) - (1.96)^2(.01) \pm \sqrt{4(1.96)^4\left(\frac{1}{3} \cdot \frac{2}{3}\right)\left(\frac{1}{3} \cdot \frac{2}{3} - .01\right) + .01(1.96)^4}}{.01} \approx 339$$

26. With $\theta = \lambda$, $\hat{\theta} = \bar{X}$ and $\sigma_{\hat{\theta}} = \sqrt{\frac{\lambda}{n}}$ so $\hat{\sigma}_{\hat{\theta}} = \sqrt{\frac{\bar{X}}{n}}$. The large sample C.I. is then

$\bar{x} \pm z_{\alpha/2} \sqrt{\frac{\bar{x}}{n}}$. We calculate $\sum x_i = 203$, so $\bar{x} = 4.06$, and a 95% interval for λ is

$$4.06 \pm 1.96 \sqrt{\frac{4.06}{50}} = 4.06 \pm .56 = (3.50, 4.62)$$

27. Note that the midpoint of the new interval is $\frac{x + \frac{z^2}{2}}{n + z^2}$, which is roughly $\frac{x + 2}{n + 4}$ with a confidence level of 95% and approximating $1.96 \approx 2$. The variance of this quantity is $\frac{np(1-p)}{(n + z^2)^2}$, or roughly $\frac{p(1-p)}{n + 4}$. Now replacing p with $\frac{x + 2}{n + 4}$, we have

$$\left(\frac{x + 2}{n + 4}\right) \pm z_{\alpha/2} \sqrt{\frac{\left(\frac{x + 2}{n + 4}\right)\left(1 - \frac{x + 2}{n + 4}\right)}{n + 4}}; \text{ For clarity, let } x^* = x + 2 \text{ and } n^* = n + 4, \text{ then}$$

$$\hat{p}^* = \frac{x^*}{n^*} \text{ and the formula reduces to } \hat{p}^* \pm z_{\alpha/2} \sqrt{\frac{\hat{p}^* \hat{q}^*}{n^*}}, \text{ the desired conclusion. For}$$

further discussion, see the Agresti article.

Section 7.3

28.

a. 1.341

d. 1.684

b. 1.753

e. 2.704

c. 1.708

Chapter 7: Statistical Intervals Based on a Single Sample

29.

a. $t_{.025,10} = 2.228$

d. $t_{.005,50} = 2.678$

b. $t_{.025,20} = 2.086$

e. $t_{.01,25} = 2.485$

c. $t_{.005,20} = 2.845$

f. $-t_{.025,5} = -2.571$

30.

a. $t_{.025,10} = 2.228$

d. $t_{.005,4} = 4.604$

b. $t_{.025,15} = 2.131$

e. $t_{.01,24} = 2.492$

c. $t_{.005,15} = 2.947$

f. $t_{.005,37} \approx 2.712$

31.

a. $t_{.05,10} = 1.812$

d. $t_{.01,4} = 3.747$

b. $t_{.05,15} = 1.753$

e. $\approx t_{.025,24} = 2.064$

c. $t_{.01,15} = 2.602$

f. $t_{.01,37} \approx 2.429$

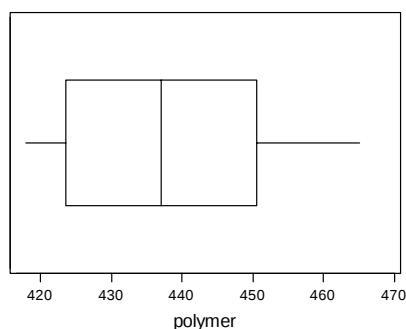
32. d.f. = $n - 1 = 7$, so the critical value for a 95% C.I. is $t_{.025,7} = 2.365$. The interval is

$$30.2 \pm \left(2.365 \left(\frac{3.1}{\sqrt{8}} \right) \right) = 30.2 \pm 2.6 = (27.6, 32.8).$$

Chapter 7: Statistical Intervals Based on a Single Sample

33.

- a. The boxplot indicates a very slight positive skew, with no outliers. The data appears to center near 438.



- b. Based on a normal probability plot, it is reasonable to assume the sample observations came from a normal distribution.
- c. With d.f. = $n - 1 = 16$, the critical value for a 95% C.I. is $t_{.025,16} = 2.120$, and the interval is $438.29 \pm (2.120) \left(\frac{15.14}{\sqrt{17}} \right) = 438.29 \pm 7.785 = (430.51, 446.08)$. Since 440 is within the interval, 440 is a plausible value for the true mean. 450, however, is not, since it lies outside the interval.

34. $n = 14$, $\bar{x} = 8.48$, $s = .79$; $t_{.05,13} = 1.771$

- a. A 95% lower confidence bound: $8.48 - 1.771 \left(\frac{.79}{\sqrt{14}} \right) = 8.48 - .37 = 8.11$. With 95% confidence, the value of the true mean proportional limit stress of all such joints lies in the interval $(8.11, \infty)$. If this interval is calculated for sample after sample, in the long run 95% of these intervals will include the true mean proportional limit stress of all such joints. We must assume that the sample observations were taken from a normally distributed population.
- b. A 95% lower prediction bound: $8.48 - 1.771(.79) \sqrt{1 + \frac{1}{14}} = 8.48 - 1.45 = 7.03$. If this bound is calculated for sample after sample, in the long run 95% of these bounds will provide a lower bound for the corresponding future values of the proportional limit stress of a single joint of this type.

Chapter 7: Statistical Intervals Based on a Single Sample

35. $n = 5, \bar{x} = 2887.6, s = .84.0; t_{.025,4} = 2.776$

a. A 95% C.I. for the mean: $2887.6 \pm (2.776)\left(\frac{.84}{\sqrt{5}}\right) \Rightarrow (2783.3, 2991.9)$

b. A 95% Prediction Interval: $2887.6 \pm 2.776(.84)\sqrt{1 + \frac{1}{5}} \Rightarrow (2632.1, 3143.1)$. The P.I. is considerably larger than the C.I., about 2.5 times larger.

36. $n = 26, \bar{x} = 370.69, s = 24.36; t_{.05,25} = 1.708$

a. A 95% upper confidence bound:

$$370.69 + (1.708)\left(\frac{24.36}{\sqrt{26}}\right) = 370.69 + 8.16 = 378.85$$

b. A 95% upper prediction bound:

$$370.69 + 1.708(24.36)\sqrt{1 + \frac{1}{26}} = 370.69 + 42.45 = 413.14$$

c. Following a similar argument as that on p. 300 of the text, we need to find the variance of

$$\begin{aligned} \bar{X} - \bar{X}_{new} : V(\bar{X} - \bar{X}_{new}) &= V(\bar{X}) + V(\bar{X}_{new}) = V(\bar{X}) + V\left(\frac{1}{2}(X_{27} + X_{28})\right) \\ &= V(\bar{X}) + V\left(\frac{1}{2}X_{27}\right) + V\left(\frac{1}{2}X_{28}\right) = V(\bar{X}) + \frac{1}{4}V(X_{27}) + \frac{1}{4}V(X_{28}) \\ &= \frac{\sigma^2}{n} + \frac{1}{4}\sigma^2 + \frac{1}{4}\sigma^2 = \sigma^2\left(\frac{1}{2} + \frac{1}{n}\right). \end{aligned}$$

We eventually arrive at $T = \frac{\bar{X} - \bar{X}_{new}}{s\sqrt{\frac{1}{2} + \frac{1}{n}}} \sim t$

distribution with $n - 1$ d.f., so the new prediction interval is $\bar{x} \pm t_{\alpha/2, n-1} \cdot s\sqrt{\frac{1}{2} + \frac{1}{n}}$. For this situation, we have

$$370.69 \pm 1.708(24.36)\sqrt{\frac{1}{2} + \frac{1}{26}} = 370.69 \pm 30.53 = (39.47, 400.53)$$

37.

a. A 95% C.I. : $.9255 \pm 2.093(.0181) = .9255 \pm .0379 \Rightarrow (.8876, .9634)$

b. A 95% P.I. : $.9255 \pm 2.093(.0809)\sqrt{1 + \frac{1}{20}} = .9255 \pm .1735 \Rightarrow (.7520, 1.0990)$

c. A tolerance interval is requested, with $k = 99$, confidence level 95%, and $n = 20$. The tolerance critical value, from Table A.6, is 3.615. The interval is $.9255 \pm 3.615(.0809) \Rightarrow (.6330, 1.2180)$.

Chapter 7: Statistical Intervals Based on a Single Sample

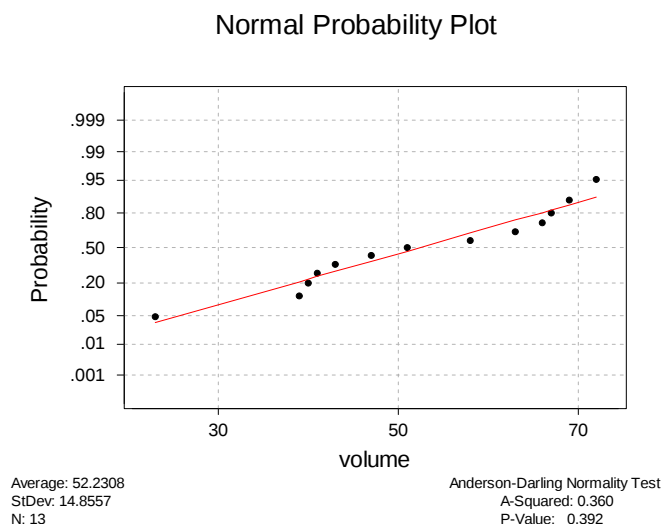
38. $N = 25$, $\bar{x} = .0635$, $s = .0065$

a. 95% P.I. : $.0635 \pm 2.064(.0065)\sqrt{1 + \frac{1}{25}} = .0635 \pm .0137 \Rightarrow (.0498, .0772)$.

b. 99% Tolerance Interval, with $k = 95$, critical value 2.972 (table A.6):
 $.0635 \pm 2.972(.0065) \Rightarrow (.0442, .0828)$.

39.

a.



Based on the above plot, generated by Minitab, it is plausible that the population distribution is normal.

b. We require a tolerance interval. (from table A6, with 95% confidence, $k = 95$, and $n=13$, the $tcv = 3.081$.)

$$\bar{x} \pm (tcv)s = 52.231 \pm 3.081(14.856) = 52.231 \pm 45.771 \Rightarrow (6.460, 98.002)$$

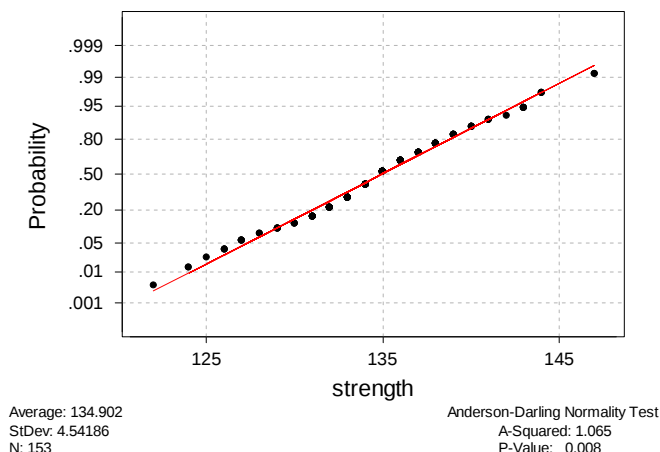
c. A prediction interval, with $t_{.025,12} = 2.179$:

$$52.231 \pm 2.179(14.856)\sqrt{1 + \frac{1}{13}} = 52.231 \pm 33.593 \Rightarrow (18.638, 85.824)$$

Chapter 7: Statistical Intervals Based on a Single Sample

40.

- a. We need to assume the samples came from a normally distributed population.
- b. A Normal Probability plot, generated by Minitab:
Normal Probability Plot



The very small p-value indicates that the population distribution from which this data was taken is most likely not normal.

- c. 95% lower prediction bound:

$$52.231 \pm 2.179(14.856)\sqrt{1 + \frac{1}{13}} = 52.231 \pm 33.593 \Rightarrow (18.638, 85.824)$$

41. The 20 d.f. row of Table A.5 shows that 1.725 captures upper tail area .05 and 1.325 captures uppertail area .10. The confidence level for each interval is 100(central area)%. For the first interval, central area = 1 – sum of tail areas = 1 – (.25 + .05) = .70, and for the second and third intervals the central areas are 1 – (.20 + .10) = .70 and 1 – (.15 + .15) = .70. Thus each interval has confidence level 70%. The width of the first interval is

$$\frac{s(.687 + 1.725)}{\sqrt{n}} = \frac{.2412s}{\sqrt{n}}, \text{ whereas the widths of the second and third intervals are } 2.185$$

and 2.128 respectively. The third interval, with symmetrically placed critical values, is the shortest, so it should be used. This will always be true for a t interval.

Section 7.2

42.

a. $\chi^2_{.1,15} = 22.307$ (.1 column, 15 d.f. row)

d. $\chi^2_{.005,25} = 46.925$

b. $\chi^2_{.1,25} = 34.381$

e. $\chi^2_{.99,25} = 11.523$ (from .99 column, 25 d.f. row)

c. $\chi^2_{.01,25} = 44.313$

f. $\chi^2_{.995,25} = 10.519$

43.

a. $\chi^2_{.05,10} = 18.307$

b. $\chi^2_{.95,10} = 3.940$

c. Since $10.987 = \chi^2_{.975,22}$ and $36.78 = \chi^2_{.025,22}$, $P(\chi^2_{.975,22} \leq \chi^2 \leq \chi^2_{.025,22}) = .95$.

d. Since $14.61 = \chi^2_{.95,25}$ and $37.65 = \chi^2_{.05,25}$, $P(\chi^2_{.95,25} \leq \chi^2 \leq \chi^2_{.05,25}) = .90$.

44. $n - 1 = 8$, $\chi^2_{.025,8} = 17.543$, $\chi^2_{.975,8} = 2.180$, so the 95% interval for σ^2 is

$$\left(\frac{8(7.90)}{17.543}, \frac{8(7.90)}{2.180} \right) = (3.60, 28.98). \text{ The 95\% interval for } \sigma \text{ is } (\sqrt{3.60}, \sqrt{28.98}) = (1.90, 5.38).$$

45. $n = 22$ implies that d.f. = $n - 1 = 21$, so the .995 and .005 columns of Table A.7 give the necessary chi-squared critical values as 8.033 and 41.399. $\sum x_i = 1701.3$ and

$\sum x_i^2 = 132,097.35$, so $s^2 = 25.368$. The interval for σ^2 is

$$\left(\frac{21(25.368)}{41.399}, \frac{21(25.368)}{8.033} \right) = (12.868, 66.317) \text{ and that for } \sigma \text{ is } (3.6, 8.1) \text{ Validity of this interval requires that fracture toughness be (at least approximately) normally distributed.}$$

46.

a. Using a normal probability plot, we ascertain that it is plausible that this sample was taken from a normal population distribution.

b. With $s = 1.579$, $n = 15$, and $\chi^2_{.05,14} = 23.685$ the 95% upper confidence bound for σ

$$\text{is } \sqrt{\frac{14(1.579)^2}{23.685}} = 1.214$$

Supplementary Exercises

47.

- a. $n = 48$, $\bar{x} = 8.079$, $s^2 = 23.7017$, and $s = 4.868$.

A 95% C.I. for μ = the true average strength is

$$\bar{x} \pm 1.96 \frac{s}{\sqrt{n}} = 8.079 \pm 1.96 \frac{4.868}{\sqrt{48}} = 8.079 \pm 1.377 = (6.702, 9.456)$$

- b. $\hat{p} = \frac{13}{48} = .2708$. A 95% C.I. is

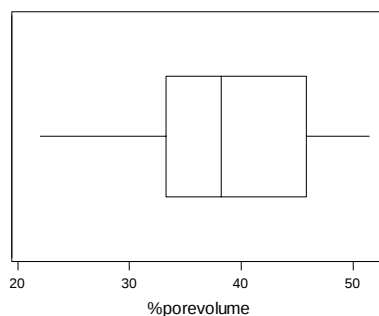
$$\frac{.2708 + \frac{1.96^2}{2(48)} \pm 1.96 \sqrt{\frac{(.2708)(.7292)}{48} + \frac{1.96^2}{4(48)^2}}}{1 + \frac{1.96^2}{48}} = \frac{.3108 \pm .1319}{1.0800} = (.166, .410)$$

48. A 98% t C.I. requires $t_{\alpha/2, n-1} = t_{.01, 8} = 2.896$. The interval is

$$188.0 \pm 2.896 \frac{7.2}{\sqrt{9}} = 188.0 \pm 7.0 = (181.0, 195.0).$$

49.

- a. There appears to be a slight positive skew in the middle half of the sample, but the lower whisker is much longer than the upper whisker. The extent of variability is rather substantial, although there are no outliers.



- b. The pattern of points in a normal probability plot is reasonably linear, so, yes, normality is plausible.

- c. $n = 18$, $\bar{x} = 38.66$, $s = 8.473$, and $t_{.01, 17} = 2.586$. The 98% confidence interval is

$$38.66 \pm 2.586 \frac{8.473}{\sqrt{18}} = 38.66 \pm 5.13 = (33.53, 43.79).$$

Chapter 7: Statistical Intervals Based on a Single Sample

50. $\bar{x}$ = the middle of the interval = $\frac{229.764 + 233.502}{2} = 231.633$. To find s we use

$$width = 2\left(t_{.025,4}\right)\left(\frac{s}{\sqrt{n}}\right), \text{ and solve for } s. \text{ Here, } n = 5, t_{.025,4} = 2.776, \text{ and width} = \text{upper}$$

$$\text{limit} - \text{lower limit} = 3.738. 3.738 = 2(2.776)\frac{s}{\sqrt{5}} \Rightarrow s = \frac{\sqrt{5}(3.738)}{2(2.776)} = 1.5055. \text{ So for}$$

a 99% C.I., $t_{.005,4} = 4.604$, and the interval is

$$231.633 \pm 4.604 \frac{1.5055}{\sqrt{5}} = 231.633 \pm 3.100 = (228.533, 234.733).$$

51.

a. $\hat{p} = \frac{136}{200} = .680 \Rightarrow$ a 90% C.I. is

$$\frac{.680 + \frac{1.645^2}{2(200)} \pm 1.645 \sqrt{\frac{(.680)(.320)}{200} + \frac{1.645^2}{4(200)^2}}}{1 + \frac{1.645^2}{200}} = \frac{.6868 \pm .0547}{1.01353} = (.624, .732)$$

b. $n = \frac{2(1.645)^2(.25) - (1.645)^2(.05)^2 \pm \sqrt{4(1.645)^4(.25)(.25 - .0025) + .05^2(1.645)^4}}{.0025}$

$$= \frac{1.3462 \pm 1.3530}{.0025} = 1079.7 \Rightarrow \text{use } n = 1080$$

c. No, it gives a 95% upper bound.

52.

a. Assuming normality, $t_{.05,15} = 1.753$, do a 95% C.I. for μ is

$$.214 \pm 1.753 \frac{.036}{\sqrt{16}} = .214 \pm .016 = (.198, .230)$$

b. A 90% upper bound for σ , with $\chi^2_{.10,15} = 1.341$, is $\sqrt{\frac{15(.036)^2}{1.341}} = \sqrt{.0145} = .120$

c. A 95% prediction interval, with $t_{.025,15} = 2.131$, is

$$.214 \pm 2.131(.036)\sqrt{1 + \frac{1}{16}} = .214 \pm .0791 = (.1349, .2931).$$

Chapter 7: Statistical Intervals Based on a Single Sample

53. With $\hat{\theta} = \frac{1}{3}(\bar{X}_1 + \bar{X}_2 + \bar{X}_3) - \bar{X}_4$, $\sigma_{\hat{\theta}}^2 = \frac{1}{9}\text{Var}(\bar{X}_1 + \bar{X}_2 + \bar{X}_3) + \text{Var}(\bar{X}_4) = \frac{1}{9}\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} + \frac{\sigma_3^2}{n_3}\right) + \frac{\sigma_4^2}{n_4}$; $\hat{\sigma}_{\hat{\theta}}$ is obtained by replacing each $\hat{\sigma}_i^2$ by s_i^2 and taking the square root. The large-sample interval for θ is then
- $$\frac{1}{3}(\bar{X}_1 + \bar{X}_2 + \bar{X}_3) - \bar{X}_4 \pm z_{\alpha/2} \sqrt{\frac{1}{9}\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} + \frac{s_3^2}{n_3}\right) + \frac{s_4^2}{n_4}}.$$
- For the given data, $\hat{\theta} = -.50$, $\hat{\sigma}_{\hat{\theta}} = .1718$, so the interval is $-.50 \pm 1.96(.1718) = (-.84, -.16)$.
54. $\hat{p} = \frac{11}{55} = .2 \Rightarrow$ a 90% C.I. is
- $$\frac{.2 + \frac{1.645^2}{2(55)} \pm 1.645 \sqrt{\frac{(.2)(.8)}{55} + \frac{1.645^2}{4(55)^2}}}{1 + \frac{1.645^2}{55}} = \frac{.2246 \pm .0887}{1.0492} = (.1295, .2986).$$
55. The specified condition is that the interval be length .2, so $n = \left[\frac{2(1.96)(.8)}{.2} \right]^2 = 245.86$, so $n = 246$ should be used.
- 56.
- A normal probability plot lends support to the assumption that pulmonary compliance is normally distributed. Note also that the lower and upper fourths are 192.3 and 228.1, so the fourth spread is 35.8, and the sample contains no outliers.
 - $t_{.025,15} = 2.131$, so the C.I. is
- $$209.75 \pm 2.131 \frac{24.156}{\sqrt{16}} = 209.75 \pm 12.87 = (196.88, 222.62).$$
- $K = 95$, $n = 16$, and the tolerance critical value is 2.903, so the 95% tolerance interval is $209.75 \pm 2.903(24.156) = 209.75 \pm 70.125 = (139.625, 279.875)$.
57. Proceeding as in Example 7.5 with T_r replacing $\sum X_i$, the C.I. for $\frac{1}{\lambda}$ is $\left(\frac{2t_r}{\chi_{1-\alpha/2, 2r}^2}, \frac{2t_r}{\chi_{\alpha/2, 2r}^2} \right)$ where $t_r = y_1 + \dots + y_r + (n-r)y_r$. In Example 6.7, $n = 20$, $r = 10$, and $t_r = 1115$. With d.f. = 20, the necessary critical values are 9.591 and 34.170, giving the interval (65.3, 232.5). This is obviously an extremely wide interval. The censored experiment provides less information about $\frac{1}{\lambda}$ than would an uncensored experiment with $n = 20$.

58.

$$\begin{aligned}
 \text{a. } P(\min(X_i) \leq \tilde{\mu} \leq \max(X_i)) &= 1 - P(\tilde{\mu} < \min(X_i) \text{ or } \max(X_i) < \tilde{\mu}) \\
 &= 1 - P(\tilde{\mu} < \min(X_i)) - P(\max(X_i) < \tilde{\mu}) \\
 &= 1 - P(\tilde{\mu} < X_1, \dots, \tilde{\mu} < X_n) - P(X_1 < \tilde{\mu}, \dots, X_n < \tilde{\mu}) \\
 &= 1 - (.5)^n - (.5)^n = 1 - 2(.5)^{n-1}, \text{ from which the confidence interval follows.}
 \end{aligned}$$

b. Since $\min(x_i) = 1.44$ and $\max(x_i) = 3.54$, the C.I. is (1.44, 3.54).

$$\begin{aligned}
 \text{c. } P(X_{(2)} \leq \tilde{\mu} \leq X_{(n-1)}) &= 1 - P(\tilde{\mu} < X_{(2)}) - P(X_{(n-1)} < \tilde{\mu}) \\
 &= 1 - P(\text{at most one } X_i \text{ is below } \tilde{\mu}) - P(\text{at most one } X_i \text{ exceeds } \tilde{\mu}) \\
 &= 1 - (.5)^n - \binom{n}{1}(.5)^1(.5)^{n-1} - (.5)^n - \binom{n}{1}(.5)^{n-1}(.5) \\
 &= 1 - 2(n+1)(.5)^n = 1 - (n+1)(.5)^{n-1}
 \end{aligned}$$

Thus the confidence coefficient is $1 - (n+1)(.5)^{n-1}$, or in another way, a $100(1 - (n+1)(.5)^{n-1})\%$ confidence interval.

59.

$$\text{a. } \int_{(\alpha/2)^{1/n}}^{(1-\alpha/2)^{1/n}} nu^{n-1} du = u^n \Big|_{(\alpha/2)^{1/n}}^{(1-\alpha/2)^{1/n}} = 1 - \frac{\alpha}{2} - \frac{\alpha}{2} = 1 - \alpha. \text{ From the probability}$$

statement, $\frac{(\alpha/2)^{1/n}}{\max(X_i)} \leq \frac{1}{\theta} \leq \frac{(1-\alpha/2)^{1/n}}{\max(X_i)}$ with probability $1 - \alpha$, so taking the

reciprocal of each endpoint and interchanging gives the C.I. $\left(\frac{\max(X_i)}{(1-\alpha/2)^{1/n}}, \frac{\max(X_i)}{(\alpha/2)^{1/n}} \right)$

for θ .

$$\text{b. } \alpha^{1/n} \leq \frac{\max(X_i)}{\theta} \leq 1 \text{ with probability } 1 - \alpha, \text{ so } 1 \leq \frac{\theta}{\max(X_i)} \leq \frac{1}{\alpha^{1/n}} \text{ with}$$

probability $1 - \alpha$, which yields the interval $\left(\max(X_i), \frac{\max(X_i)}{\alpha^{1/n}} \right)$.

c. It is easily verified that the interval of **b** is shorter – draw a graph of $f_U(u)$ and verify that the shortest interval which captures area $1 - \alpha$ under the curve is the rightmost such interval, which leads to the C.I. of **b**. With $\alpha = .05$, $n = 5$, $\max(x_i) = 4.2$; this yields (4.2, 7.65).

Chapter 7: Statistical Intervals Based on a Single Sample

60. The length of the interval is $(z_\gamma + z_{\alpha-\gamma}) \frac{s}{\sqrt{n}}$, which is minimized when $z_\gamma + z_{\alpha-\gamma}$ is minimized, i.e. when $\Phi^{-1}(1-\gamma) + \Phi^{-1}(1-\alpha+\gamma)$ is minimized. Taking $\frac{d}{d\gamma}$ and equating to 0 yields $\frac{1}{\Phi(1-\gamma)} = \frac{1}{\Phi(1-\alpha+\gamma)}$ where $\Phi(\bullet)$ is the standard normal p.d.f., whence $\gamma = \frac{\alpha}{2}$.

61. $\tilde{x} = 76.2$, the lower and upper fourths are 73.5 and 79.7, respectively, and $f_s = 6.2$. The robust interval is $76.2 \pm (1.93) \left(\frac{6.2}{\sqrt{22}} \right) = 76.2 \pm 2.6 = (73.6, 78.8)$.
 $\bar{x} = 77.33$, $s = 5.037$, and $t_{.025, 21} = 2.080$, so the t interval is $77.33 \pm (2.080) \left(\frac{5.037}{\sqrt{22}} \right) = 77.33 \pm 2.23 = (75.1, 79.6)$. The t interval is centered at $\bar{x}$, which is pulled out to the right of $\tilde{x}$ by the single mild outlier 93.7; the interval widths are comparable.

62.

- a. Since $2\lambda \Sigma X_i$ has a chi-squared distribution with $2n$ d.f. and the area under this chi-squared curve to the right of $\chi_{.95, 2n}^2$ is .95, $P(\chi_{.95, 2n}^2 < 2\lambda \Sigma X_i) = .95$. This implies that $\frac{\chi_{.95, 2n}^2}{2\Sigma X_i}$ is a lower confidence bound for λ with confidence coefficient 95%. Table A.7 gives the chi-squared critical value for 20 d.f. as 10.851, so the bound is $\frac{10.851}{2(550.87)} = .0098$. We can be 95% confident that λ exceeds .0098.

- b. Arguing as in a, $P(2\lambda \Sigma X_i < \chi_{.05, 2n}^2) = .95$. The following inequalities are equivalent to the one in parentheses:

$$\lambda < \frac{\chi_{.05, 2n}^2}{2\Sigma X_i} \Rightarrow -\lambda t < \frac{-t\chi_{.05, 2n}^2}{2\Sigma X_i} \Rightarrow e^{-\lambda t} < \exp\left[\frac{-t\chi_{.05, 2n}^2}{2\Sigma X_i}\right].$$

Replacing the ΣX_i by Σx_i in the expression on the right hand side of the last inequality gives a 95% lower confidence bound for $e^{-\lambda t}$. Substituting $t = 100$, $\chi_{.05, 20}^2 = 31.410$ and $\Sigma x_i = 550.87$ gives .058 as the lower bound for the probability that time until breakdown exceeds 100 minutes.